

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.

THIS PAGE BLANK (USPTO)

PRI 22-FEB-2000

Alignment Scores:	
Pred. No.:	1.26e-197
length:	2820

Score:	2386.00	Matches:	460
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-001-254-16 (1-460) x AK000528 (1-2820)

QY	1	MeAaenlyspProiletrProserThThyValArgCysleuSnValGlyleuIleArg	20
DB	30	ATGAACAAACCATTAAACCATCAACATATGTGGCTGCTCAATGTGGACATTAATAG	89
QY	21	LysLeuSerAerPheIleaspProGlnGlyTrpIlySylsleuAlaValaIleLys	40
DB	90	AAGCTGCAGATTTTATTGATCTCAAGAAAGATGAAGAAGTTAGCTAGCTATTAAA	149
QY	41	LysProSerGlyAerAerAerGlyAerGlnPheHisIleArgArgPheGlyAlaLeuLeu	60
DB	150	AAACCATCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	209
QY	61	GlnTrpGlyLysSerProThSerGlyLeuLeuPheAerPTrpGlyThrThraEncyThr	80
DB	210	CAAACTGGAAAAAGTCCACCTTCGATTTACTGTTTACCTGGGGCACCAATTCACA	269
QY	81	ValGlyAerLeuValAerLeuLeuIleGlnAsnGlnPheAerAerProAerSerLeuLeu	100
DB	270	GTGGTGATCTTTGGATCTTTGATCCAAATGATTTTCTCTCGGAGATCTTTGG	329
QY	101	LeuProAerPAalAalProLySerThAlaAsnThreupProSerLySGlyAlaIleThrVal	120
DB	330	CTCCCAAGATGCTGTTCCCAAACTGCTAAACACTTACTCTTAAAGACTTAACAGTT	389
QY	121	GlnGlnLySGlnMetProPheCysAerAerAerGlyLeuMetThrProValGlnAsn	140
DB	390	CAGCAAAAAACAGATGCTCTTCTGTGACAAAGACAGACATTTGATGACACTGGACAGAT	449
QY	141	LeuGlnGlnSerTyMetProProAerSerSerProGlnSylsSerLeuGlyVal	160
DB	450	CTTGAACAAAGCTATATGCCACCTGACTCCCAAGTCCAAATAAATTAAGTTTGAAGTT	539
QY	161	SerAerTrpAerPheHisSerPheSerPheTyGlyLeuLysAsnValThraAsnAerPhe	180
DB	510	AATGATACACGTTTTCAAGTTTTTCAATTTTATGAATTAAGAAGATGACAAATTAATCTT	569
QY	181	AerGlyAerProIleSerValGlyGlyAsnLysMetGlyGlyGlyPheGlyValVal	200
DB	570	GATGAACACACCATTTCTGTTGGTGATTAATAATGGGAAGAGAGATTTGGATGTTGA	629
QY	201	TyTrpSGlyTyValAerAerThThValAlaValLysLysLysLeuAlaIleMetValAer	220
DB	630	TATTAAGGCTACGTAAATTAACACACTGTGGCACTGAAGACTTGCAGATAGCTTGAC	689
QY	221	IleThrTrpGlnGlyLeuLysGlnIlePheAerPTrpIleLysValMetAlaLysCys	240
DB	690	ATTACTACTGAAGAACTGAACACAGACTTGTATCAAAATAAAGATTAAGCAAAAGTGT	749
QY	241	GlnHisGlyAerLeuValGlyLeuLeuGlyPheSerSerAerAerAerAerAerAerAerAer	260
DB	750	CAACATGAAAACTTAATGTAACCTAATGTTTCCAAAGTATGAGATGAGACCTTCTCTTA	819
QY	261	ValTyValTyMetProAerGlySerLeuLeuAerAerAerAerAerAerAerAerAerAer	280
DB	810	GATATGTTTACATGCTCAATGTTTCAATGTTTCAATGTTTCAATGTTTCAATGTTTCAAT	869
QY	281	ProProLeuSerTrpHisMetArgCysLysIleAlaGlnGlyAlaAlaAsnGlyIleAsn	300
DB	870	CCACCACTTTCTTGGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT	929
QY	301	PheLeuHisGlyAerHisHisIleHisArgAerAerAerAerAerAerAerAerAerAer	320
DB	930	TTTCTACATGAAATCATCATATTTCTATAGATATTTAAAGTGCMAATATCTTACTGAT	989
QY	321	GlnAlaPheThrAlaLysIleSerAerPheGlyLeuAlaArgAerGlyLysPheAla	340

THIS PAGE BLANK (USPTO)

DB 990 GAAGTTTACTGCTAAATATCTGACTTGTGCGCTTGACGGCTTCTGAGAAATTTGCC 1049
Qy 341 GlnThrValMetThrSerArgIleValGlyThrThralaTyMetAlaProGluAlaLeu 360
Db 1050 CAGACAGTCATGACTAGCAGAAATGTGGGAAACAAGCTTATATGACCAAGAGCTTTG 1109
Qy 361 ArgGlyGluIleThrProLysSerAspIleTyrSerPheGlyValValLeuLeuGluIle 380
Db 1110 CGTGAGAAATTAACCCCAAAATGATATTAACAGCTTGTGGTGTATTAAGTAAATA 1169
Qy 381 IleThrGlyLeuProAlaValAspGluHisArgGluProGluLeuLeuAspIleLys 400
Db 1170 ATACTGACTTCCAGCTGTGATGATACACCGTGAACCTCACTTATCTGATTTAA 1229
Qy 401 GluGluIleGluAspGluGluLysThrIleGluAspTyrIleAspLysLysMetAsnAsp 420
Db 1230 GAGCAATTTGAAGTGAAGAAAGACAAATGATGATATTTGATTAAGATGATGAT 1289
Qy 421 AlaAspSerThrSerValGluAlaMetTyrSerValAlaSerGlnCysLeuHisGluLys 440
Db 1290 GCTGATTCACCTCAGTGAAGCTATGATCTGCTGCTAGCAATGCTCATGAGAAAG 1349
Qy 441 LysAsnLysArgProAspIleLysLysValGlnGluLeuGluGluMetThrAlaSer 460
Db 1350 AAAAATAAGAGCCAGACATTAGAGAGTTCAACAGCTGCTGCAGAGATGACAGCTTCT 1409

RESULT 4
AX196260 1383 bp DNA linear PAT 28-AUG-2001
LOCUS AX196260
DEFINITION Sequence 2 from Patent WO0151641.
ACCESSION AX196260
VERSION AX196260.1 GI:15386462
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Weeche, H. and Li, S.
TITLE Irak-4: compositions and methods of use
JOURNAL Patent: WO 0151641-A 2 19-JUL-2001;
Tularik Inc. (US)
FEATURES
source Location/Qualifiers
1..1383
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="human IL-1 receptor-associated kinase 4 (IRAK-4)
cDNA"
1..1383
/note="human IRAK-4"
/codon_start=1
/protein_id="CAC60090.1"
/db_xref="GI:15386463"
/translation="MNKPIPTSTYVRCINVLIRKLSDFIDPQEMKKLAVAIKKPSG
DDRNVPHIRREALLQTKSPSELDDMGTTNCTADVLIDVLIQNEFFAPALILP
DAVPTANTLPSKEAIIYQCKMPKCDKRLIMPTVLEBSYPPSSSEMSISLY
SDTRHSFSFELKAVTNFDERPI SVGKMKGBGFEVTKGVNNTVAVKLAAM
VDITTEELKQDQDEIKYAKACQHNELVELGSSGDGDLVYVYMGSLDRISC
LDGTPPLSWHMKCKIAOGANGINFLHNHHRDIKSNAILLDEAFPAKISDFGLAR
ASEKPAQVMTSRIVGTAYVMAPEALRGELTPKSDIYSGVVLLEITGLPAVNEHRE
PQLLDIKKEIEDEKTIEDYIDKMNADSTSVAMYSVASCHEKKRNPDIKKY
QQLQEMTAS"

BASE COUNT 463 a 244 c 283 g 393 t
ORIGIN
Alignment Scores:
Pred. No.: 1,11e-197 Length: 1383
Score: 2382.00 Matches: 459
Percent Similarity: 99.788 Conservative: 0
Best Local Similarity: 99.788 Mismatches: 0
Query Match: 99.83% Indels: 0

Qy 341 GlnThrValMetThrSerArgIleValGlyThrThralaTyMetAlaProGluAlaLeu 360

DB: 6 Gaps: 0
US-10-001-254-16 (1-460) x AX196260 (1-1383)
Qy 1 MetAsnLysProIleThrProSerThrTyrValAlaArgCysLeuAsnValGlyLeuIleArg 20
Db 1 ATGACAAACCCATTAACACATCAATCAATATGTCGCCGTGCTCAATGTTGACATTAATAG 60
Qy 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPheLysLeuAlaValAlaIleLys 40
Db 61 AACCTTCAGATTTTATTAATCTCTCAAGAGATGGAAACCTTGTGCTTACTTAA 120
Qy 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgArgPheGluAlaLeuLeu 60
Db 121 AACCATCTGGTATATATGATACATCACTTTCACATTAAGAGATTTGACATTAAT 180
Qy 61 GlnThrGlyLysSerProThrSerGluLeuLeuPheAspTyrGlyThrThrAsnCysThr 80
Db 181 CAAACTGAAAAAGTCCCACTTCTGAAATTAAGTTGACTGGGGCCACCAAAATTCACACA 240
Qy 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGluPhePheAlaProAlaSerLeuLeu 100
Db 241 GCTGGTATCTGTGTGATCTTTGATCCAAATGAATTTTGTCTGCGAGTCTTTTG 300
Qy 101 LeuProAspAlaValProLysThrAlaAsnThrLeuProSerLysGluAlaIleThrVal 120
Db 301 CTCCCAAGATGCTTCTCCCAAACTGCTAATACCTTCTTAAAGAACTATACAGATT 360
Qy 121 GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140
Db 361 CAGCAAAAACAGATGCTTCTGTGACAAAGACAGACATTTGATGACACCTGTCAGAAAT 420
Qy 141 LeuGluGlnSerTyrMetProProAspSerSerProGluAsnLysSerLeuGluVal 160
Db 421 CTGAAACAAAGCTATATGACCACTGCTCAAGTCCAGAAATTAAGTTAGAAAGTT 480
Qy 161 SerAspThrArgPheHisSerPheSerPheTyrGluLeuLysAsnValThrAsnAspPhe 180
Db 481 AGTGATACAGCTTTTTCACAGTTTTCATTTTGAATTAAGATGTCACAAATACCTT 540
Qy 181 AspGluArgProIleSerValGlyGlyAsnLysMetGlyGluGlyGlyPheGlyValVal 200
Db 541 GATGAACGACCCCACTTCTGTGTGCTGAATTAAGATGAGAGAGAGATTTGAGAGTTGTA 600
Qy 201 TyrLeuGlyTyrValAlaAsnAsnThrThrValAlaValLysLysLeuAlaMetValAsp 220
Db 601 TATTAAGGCTACGTAATAATACACACTGTGAGAGAGAGATTTGACCAATGTTGAC 660
Qy 221 IleThrThrGluGluLeuLysGlnPheAspGlnGluIleLysValMetAlaLysCys 240
Db 661 ATTACTACGTAAGAACTGAAGACAGCTTGTGATCAAGAAATTAAGTATGCAAGTGT 720
Qy 241 GlnHisGluAsnLeuValGluLeuLeuGlyPheSerSerAspGlyAspAspLeuLysLeu 260
Db 721 CAACATGAAGAACTTACTAAGACTACTGTTCTCAAGATGAGAGATGACCTGCTGCTTA 780
Qy 261 ValTyrValTyrMetProAsnGlySerLeuLeuAspArgLysSerCysLeuAspGlyThr 280
Db 781 GATATAGTTTACATGCTTATGATGCTTATGCTAGACAGACTCTTCTGCTGATGATGACT 840
Qy 281 ProProLeuSerThrHisMetArgCysLysIleAlaGlnGlyAlaAlaAsnGlyIleAsn 300
Db 841 CACCACTTTCTTGGCAGATGATGACCAAGATTTGCTCAGGGTGAGCTATATGATCATAT 900
Qy 301 PheLeuHisGluAsnHisIleHisIleAspArgIleLysSerAlaAsnIleLeuLeuAsp 320
Db 901 TTTCTACATGAATAATCATATATTCATAGATATTTAAAGTCAAAATATCTTACTGAT 960
Qy 321 GluAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGluLysPheAla 340
Db 961 GAAAGCTTTTACTGCTAAATATCTGACTTGGCTTGGCAGCGGCTTCTGAGAAAGTTTGC 1020
Qy 341 GlnThrValMetThrSerArgIleValGlyThrThralaTyMetAlaProGluAlaLeu 360

THIS PAGE BLANK (USPTO)

DNA of sea ID#5

RESULT 6
AK000528
LOCUS Home sapiens cDNA FLJ20521 fls, clone KAT10395.
DEFINITION
AK000528
ACCESSION
AK000528.1 GI:7020683
VERSION
oligo capping, fls (full insert sequence).
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases) Mekanadeh, D., Kumagai, A., Itakura, S., Yamazaki, M., Tashiro, H.,
Ota, T., Suzuki, Y., Oabayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project
JOURNAL
Unpublished
AUTHORS
2 (bases 1 to 2820)
Sugano, S., Suzuki, Y., Ota, T., Oabayashi, M., Nishi, T., Isogai, T.,
Shibahara, T., Tanaka, T. and Nakamura, Y.
Direct Submission
TITLE
Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
Minato-ku Tokyo 108-8639, Japan (E-mail:cdat@ims.u-tokyo.ac.jp,
Tel:81-3-5449-5286, Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5' & 3' end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source
1..2820
Location/Qualifiers
/mol_type="mRNA"
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAT10395"
/cell_line="KATO III"
/cell_type="signal-ring cell carcinoma"
/clone_1ib="KAT"
/note="Cloning vector pME18SFL3"
30..1412
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91232.1"
/db_xref="GI:7020684"
/translation="MNKPIPTSTYVRCINVGIRKLSDFIDPQEGMKLAVAIKPPSG
DDRVNPFHRRFALLOTGKSPSELFPDMCTNCTVGDIVDLIQNEFPAPASLLP
DAVPTANTLPSKEAITVOOKMPFCDKDTLMTFVQLEOSYNEPDSSEPKSLAV
SDTRHSFSFEYELKAVTNFDERPISVGNKRGEGFVYVYVNTTAVAKLALAM
VDTTEELKQDFQDEIKVMACQEHNLVELLFSSDGDCLVYVMPGSLDLRLSC
LDGTPLSWHRCKIAOGANGINFLHSHHHRDIKSNILILDEFTAKISDFGLAR
ASEKFAQVTMTRIVGTAYVAPALRGELITPKSDIYSGVLLIETGLPAVDEHRE
POLLDIKEIEDEKTIEDYIDKKNADSTSVAMSVASQCHEKKKRPDIKY
QQLDQMTAS"
BASE COUNT 940 a 534 c 579 g 767 t
ORIGIN
Query Match 100.0%; Score 294; DB 9; Length 2820;
Best Local Similarity 100.0%; Pred. No. 1,1e-64;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 174 AATCAGTTTCACATAGAGATTTGAAGATTAATCTCAAACTGGAAGAAAGTCCACTTCT 293
Oy 181 GAATTACTGTTTACACGGGGCACCACAAATTCAGAGTTGGATGTTGGATCTTTTG 240
Db 234 GAATTACTGTTTACACGGGGCACCACAAATTCAGAGTTGGATGTTGGATCTTTTG 293
Oy 241 ATCCAAATGAATTTTGTCTCTCGAGTCCTTTTCTCCAGATCTGCTCC 294
Db 294 ATCCAAATGAATTTTGTCTCTCGAGTCCTTTTCTCCAGATCTGCTCC 347
RESULT 7
AX196260
LOCUS Sequence 2 from Patent WO0151641.
DEFINITION
AX196260
ACCESSION
AX196260.1 GI:15386462
VERSION
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 Wesche, H. and Li, S.,
Irak-4: compositions and methods of use
Patent: WO 0151641-A 2 19-JUL-2001;
Tularik Inc. (US)
JOURNAL
TITLE
Location/Qualifiers
1..1383
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="human IL-1 receptor-associated kinase 4 (IRAK-4)
cDNA"
CDS
1..1383
/note="human IRAK-4"
/codon_start=1
/protein_id="CAC60090.1"
/db_xref="GI:15386463"
/translation="MNKPIPTSTYVRCINVGIRKLSDFIDPQEGMKLAVAIKPPSG
DDRVNPFHRRFALLOTGKSPSELFPDMCTNCTVGDIVDLIQNEFPAPASLLP
DAVPTANTLPSKEAITVOOKMPFCDKDTLMTFVQLEOSYNEPDSSEPKSLAV
SDTRHSFSFEYELKAVTNFDERPISVGNKRGEGFVYVYVNTTAVAKLALAM
VDTTEELKQDFQDEIKVMACQEHNLVELLFSSDGDCLVYVMPGSLDLRLSC
LDGTPLSWHRCKIAOGANGINFLHSHHHRDIKSNILILDEFTAKISDFGLAR
ASEKFAQVTMTRIVGTAYVAPALRGELITPKSDIYSGVLLIETGLPAVDEHRE
POLLDIKEIEDEKTIEDYIDKKNADSTSVAMSVASQCHEKKKRPDIKY
QQLDQMTAS"
BASE COUNT 463 a 244 c 283 g 393 t
ORIGIN
Query Match 99.5%; Score 292.4; DB 6; Length 1383;
Best Local Similarity 99.7%; Pred. No. 3,1e-64;
Matches 293; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

THIS PAGE BLANK (USPTO)

CDS

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
50..1432
/note="unnamed protein product"

/codon_start=1
/protein_id="CAD37285.1"

/translation="MNKPIPTSPYVCLNGLIRKLSDFIDPQEWKLAVALIKKPSG
DDRNOFHILRFEALLOTGKSPSELFPDMGTNCTVGLVDLLIONEFAPASLLP
DAVPTANTLPSKEALITVOOKMPCDRTMTPVONLEOSYMPDSSPENSLAV
SDTRHRSFSEYELKNTNNPDERPIVSGNKGEGGFGYVYKGYANTTVAVKLAAM
VDITTEELKQDFQEIKNVAKCOHENLVLELGSDDDLCLVYVMPGSLDLRLSC
LDGTPPLSWHMRCKIAOGANGINFLHNNHHRDIKSNILDEAFYAKISDFGLAR
ASKEPAQVMTSRIVGTAYMAPEALRGEITPKSDIYSGVVLLEITIGLPAVDHRE
POLLIDIKERIEDEKTIEDYIDKKNADSTSVAMYSGASQCRHEKNSPDICKV
HOLLQEMTAS"

BASE COUNT 912 a 547 c 586 g 772 t
ORIGIN

Alignment Scores:

Pred. No.: 1,71e-60 Length: 2817
Score: 517.00 Matches: 98
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-001-254-6 (1-98) x AX431318 (1-2817)

QY 1 ThrryValArgCysLeuAnValGlyLeuIleArgLysLeuSerAspPheIleAspPro 20
Db 74 ACATATGGCGCTGCTCAATGTTGACTAATTAAGACCTGTCAGATTATTAATGATCCT 133
QY 21 GINGLUGLYTTPlySLysLeuAlaValAlaIleLysPProSerGlyAspAspArgTyr 40
Db 134 CAAGAAGATGAGAGAGAGTACCTGATTAATAAACAATCTGGTGATGATGATAC 193
QY 41 AaNgInPheHisIleArgArgPheGluAlaLeuLeuGlnThrGlyLysSerProThrSer 60
Db 194 AATCAGTTTCACATAAGAGATTGTAAGCATTAATCAAACTGGAAGAAAGTCCCACTTCT 253
QY 61 GluLeuLeuPheAspTTPGlyThrThrAsnCysThrValGlyAspLeuValAspLeuLeu 80
Db 254 GAATTACTGTTTGACTGGGGCACCAAAATGACAGCTGGTGTATCTTGTGATCTTTTG 313
QY 81 ILeGInaNgInuPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
Db 314 ATCCAAATGAAATTTTTCCTGCTGCGAGTCTTTCCTCCAGATGCTGTTCC 367

RESULT 6
AF155118 2817 bp mRNA linear PRI 05-JAN-2000
LOCUS Homo sapiens putative protein kinase NY-REN-64 antigen mRNA,
DEFINITION complete cds.
ACCESSION AF155118
VERSION AF155118.1 GI:5360130
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2817)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Scanlan,M.J., Jordan,J.D., Williamson,B., Stockert,E., Bander,N.H.,
Jongeneel,V., Gure,A.O., Jager,D., Jager,E., Knuth,A., Chen,Y.-T.
and Old,L.J.

TITLE Antigen recognized by autologous antibody in patients with
renal-cell carcinoma

JOURNAL Int. J. Cancer 83 (4), 456-464 (1999)

MEDLINE 99438124

REFERENCE 2 (bases 1 to 2817)

AUTHORS Scanlan,M.J., Jordan,J.D., Williamson,B., Stockert,E., Bander,N.H.,

Jongeneel,V., Gure,A.O., Jager,D., Jager,E., Knuth,A., Chen,Y.-T.
and Old,L.J.
Direct Submission
Submitted (28-MAY-1999) Ludwig Institute, Sloan-Kettering
Institute, 1275 York Ave, New York, NY 10021, USA

FEATURES
Location/Qualifiers
1..2817
/organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="1973/10-4"
/cell_type="renal cell carcinoma"
50..1432
/codon_start=1
/product="putative protein kinase NY-REN-64 antigen"

CDS

/translation="MNKPIPTSPYVCLNGLIRKLSDFIDPQEWKLAVALIKKPSG
DDRNOFHILRFEALLOTGKSPSELFPDMGTNCTVGLVDLLIONEFAPASLLP
DAVPTANTLPSKEALITVOOKMPCDRTMTPVONLEOSYMPDSSPENSLAV
SDTRHRSFSEYELKNTNNPDERPIVSGNKGEGGFGYVYKGYANTTVAVKLAAM
VDITTEELKQDFQEIKNVAKCOHENLVLELGSDDDLCLVYVMPGSLDLRLSC
LDGTPPLSWHMRCKIAOGANGINFLHNNHHRDIKSNILDEAFYAKISDFGLAR
ASKEPAQVMTSRIVGTAYMAPEALRGEITPKSDIYSGVVLLEITIGLPAVDHRE
POLLIDIKERIEDEKTIEDYIDKKNADSTSVAMYSGASQCRHEKNSPDICKV
HOLLQEMTAS"

BASE COUNT 912 a 547 c 586 g 772 t
ORIGIN

Alignment Scores:

Pred. No.: 1,71e-60 Length: 2817
Score: 517.00 Matches: 98
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-001-254-6 (1-98) x AF155118 (1-2817)

QY 1 ThrryValArgCysLeuAnValGlyLeuIleArgLysLeuSerAspPheIleAspPro 20
Db 74 ACATATGGCGCTGCTCAATGTTGACTAATTAAGACCTGTCAGATTATTAATGATCCT 133
QY 21 GINGLUGLYTTPlySLysLeuAlaValAlaIleLysPProSerGlyAspAspArgTyr 40
Db 134 CAAGAAGATGAGAGAGTACCTGATTAATAAACAATCTGGTGATGATGATAC 193
QY 41 AaNgInPheHisIleArgArgPheGluAlaLeuLeuGlnThrGlyLysSerProThrSer 60
Db 194 AATCAGTTTCACATAAGAGATTGTAAGCATTAATCAAACTGGAAGAAAGTCCCACTTCT 253
QY 61 GluLeuLeuPheAspTTPGlyThrThrAsnCysThrValGlyAspLeuValAspLeuLeu 90
Db 254 GAATTACTGTTTGACTGGGGCACCAAAATGACAGCTGGTGTATCTTGTGATCTTTTG 313
QY 81 ILeGInaNgInuPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
Db 314 ATCCAAATGAAATTTTTCCTGCTGCGAGTCTTTCCTCCAGATGCTGTTCC 367

DNA encoding 2817 bp

RESULT 7
AK000528 2820 bp mRNA linear PRI 22-FEB-2000
LOCUS Homo sapiens cDNA FLJ20521 t1s, clone KAT10395.
DEFINITION
ACCESSION AK000528
VERSION AK000528.1 GI:7020683
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2817)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
Ota,T., Suzuki,Y., Oabayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,

THIS PAGE BLANK (USPTO)

TITLE Nakamura, Y., Isogai, T. and Sugano, S.
JOURNAL NEDO human cDNA sequencing project
REFERENCE 2 (bases 1 to 2820)
AUTHORS Sugano, S., Suzuki, Y., Ota, T., Ohayashi, M., Nishi, T., Isogai, T., Shihahara, T., Tanaka, T. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' - & 3' - end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
FEATURES
source Location/Qualifiers
1..2820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KAT10395"
/cell_line="KATO III"
/cell_type="signet-ring cell carcinoma"
/clone_id="KAT"
/note="Cloning vector pME18SFL3"
30..1412
/note="unnamed protein product"
/protein_id="BA91232.1"
/cdon_start=1
/db_xref="GI:7020684"
/translation="MNKPTTPSTYVRCINGLIRKLSDFIDPOEGMKLAVALIKKPSG DRRNQPHIRFEALQYKSPSELPDMGTCTAGDVLVDLIONEFFAPASLLP DAVPTANTLPSKEAITVQOKMPFCDKDRITLMPVONLEOSYMPDSSSPNKSLFV SDTRPHFSFYLKNTNPNFDERPISVGNMGSGGFGVYKGVNNTVAVKKLAM VDTTEELKQOFDOEIKVMACQENHLEHNNHHRDIKSANILLDEAFKISDFELAR LDGTPPLSMHMRCKIAGANGINLEHNNHHRDIKSANILLDEAFKISDFELAR ASERFQATVWTSRIQVTTATYAPALAGETTPKSDIYSPVLEIITGCPAYDREHRE POLLDIKKEIEDEKTEIDYIDKMDADSTSEVAMYSVASOCHHEKKNRPDIKY QOLLEEMTAS"
BASE COUNT 940 a 534 c 579 g 767 t
ORIGIN
Alignment Scores:
Pred. No.: 1,71e-60 Length: 2820
Score: 517.00 Matches: 98
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-001-254-6 (1-98) x AK000528 (1-2820)
QY 1 ThTyrValaArgCysLeuAsnValGlyLeuIleArgIysLeuSerAspPheIleAspPro 20
Db ACATATGCGCGCTGCTCAATGCTGACTAATTGGAAGCTGTCAGATTTTATGATCCT 113
QY 21 GlnGluGlyTTPlySLysLeuAlaValAlaIleLysIysProSerGIYAspAspArgTyr 40
Db 114 CAAAGAGATGGAAGAGATGCTGCTGACTATTAAAAAACATCTGGGATGATAGATAC 173
QY 41 AsnGlnPheHisIleArgArgPheGlnAlaLeuLeuGlnInnrgIYlySerProThrsSer 60
Db 174 AATCAGTTTCACATTAAGAGATTGAACATTACTTCAACTGGAAGAAAAGTCCACTTCT 233
QY 61 GluLeuLeuPheAspTrrGIYThrThrsAsnCysThrValGlyAspLeuValAspLeuLeu 80
Db 234 GAATTACTGTTGATGGGACCAACAAATTGCAAGTTGATCTTGTGATCTTTTG 293
QY 81 IleGlnAsnGluPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
|||||

Db 294 ATCCAAATGAAATTTTTCCTCCGCGAGCTTTTGTCTCCAGATGCTGTCC 347
RESULT 8
AX196260
LOCUS 1383 bp DNA linear PAT 28-AUG-2001
DEFINITION Sequence 2 from Patent WO0151641.
ACCESSION AX196260
VERSION AX196260.1 GI:15386462
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Mesche, H. and Li, S.
TITLE Irak-4: compositions and methods of use
AUTHORS Patent: WO 0151641-A 2 19-JUL-2001;
JOURNAL Tularik Inc. (US)
FEATURES
source Location/Qualifiers
1..1383
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="human IL-1 receptor-associated kinase 4 (IRAK-4) cDNA"
CDNA
1..1383
/note="human IRAK-4"
/cdon_start=1
/protein_id="CAC60090.1"
/db_xref="GI:15386463"
/translation="MNKPTTPSTYVRCINGLIRKLSDFIDPOEGMKLAVALIKKPSG DRRNQPHIRFEALQYKSPSELPDMGTCTAGDVLVDLIONEFFAPASLLP DAVPTANTLPSKEAITVQOKMPFCDKDRITLMPVONLEOSYMPDSSSPNKSLFV SDTRPHFSFYLKNTNPNFDERPISVGNMGSGGFGVYKGVNNTVAVKKLAM VDTTEELKQOFDOEIKVMACQENHLEHNNHHRDIKSANILLDEAFKISDFELAR LDGTPPLSMHMRCKIAGANGINLEHNNHHRDIKSANILLDEAFKISDFELAR ASERFQATVWTSRIQVTTATYAPALAGETTPKSDIYSPVLEIITGCPAYDREHRE POLLDIKKEIEDEKTEIDYIDKMDADSTSEVAMYSVASOCHHEKKNRPDIKY QOLLEEMTAS"
BASE COUNT 463 a 244 c 283 g 393 t
ORIGIN
Alignment Scores:
Pred. No.: 2,56e-60 Length: 1383
Score: 513.00 Matches: 97
Percent Similarity: 98.98% Conservative: 0
Best Local Similarity: 98.98% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: Gaps: 0
US-10-001-254-6 (1-98) x AX196260 (1-1383)
QY 1 ThTyrValaArgCysLeuAsnValGlyLeuIleArgIysLeuSerAspPheIleAspPro 20
Db 25 ACATATGCGCGCTGCTCAATGCTGACTAATTGGAAGCTGTCAGATTTTATGATCCT 84
QY 21 GlnGluGlyTTPlySLysLeuAlaValAlaIleLysIysProSerGIYAspAspArgTyr 40
Db 85 CAAAGAGATGGAAGAGATGCTGCTGACTATTAAAAAACATCTGGGATGATAGATAC 144
QY 41 AsnGlnPheHisIleArgArgPheGlnAlaLeuLeuGlnInnrgIYlySerProThrsSer 60
Db 145 AATCAGTTTCACATTAAGAGATTGAACATTACTTCAACTGGAAGAAAAGTCCACTTCT 204
QY 61 GluLeuLeuPheAspTrrGIYThrThrsAsnCysThrValGlyAspLeuValAspLeuLeu 80
Db 205 GAATTACTGTTGATGGGACCAACAAATTGCAAGTTGATCTTGTGATCTTTTG 264
QY 81 IleGlnAsnGluPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
Db 265 ATCCAAATGAAATTTTTCCTCCGCGAGCTTTTGTCTCCAGATGCTGTCC 318
RESULT 9

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 18, 2004, 01:11:43 ; Search time 576.135 Seconds
(without alignments)
2488.940 Million cell updates/sec

Title: US-10-001-254-26
Perfect score: 323
Sequence: 1 MNKPRTPTVRCINVLINR.....KKPSGDRYVQHFIRCCSQN 59

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Rgapop 6.0 , Rgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=x1h
-O=/c/nt2/USPTO.spool/US10001254/rnatc.16012004.152424.19723/app.query.fasta_1.1109
-DB=EST -QMT=fastcap -SUFFIX=rest -MINMATCH=0.1 -LOOPCP=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US10001254.@CGN_1_1.4382.@rnatc.16012004.152424.19723 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: EST:
1: em_estda:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_tod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	323	100.0	284	9	AA114228
2	296	91.6	576	10	BF238344
3	290	89.8	540	9	AL699213
4	290	89.8	719	10	BR696981
5	290	89.8	811	10	BG16438
6	290	89.8	859	10	BG164491
7	284	87.9	313	12	BM151935
8	282	87.3	265	9	AM436511
9	274	84.8	402	10	BE482619
10	271	83.9	453	10	BB860349
11	271	83.9	503	14	CA538859
12	271	83.9	507	10	BB866698
13	271	83.9	524	9	AM106160
14	271	83.9	575	4	BX522921
15	271	83.9	598	13	BQ552228
16	271	83.9	610	10	BB660378
17	271	83.9	637	10	BB613447
18	271	83.9	638	14	BY721552
19	271	83.9	663	14	BY726858
20	271	83.9	676	10	BB613167
21	271	83.9	1161	11	AK020397
22	271	83.9	2481	11	AK028837
23	271	83.9	2810	11	AK028028
24	269	83.3	858	10	BF687921
25	197	61.0	629	9	AL647125
26	192	59.4	520	10	BE132064
27	190	59.4	664	14	BQ359662
28	190	58.8	852	13	BQ209111
29	186	57.6	600	9	AJ453616
30	186	57.6	670	9	AJ447581
31	159	49.2	541	10	BG691069
32	135	41.8	632	14	CA365604
33	125	38.7	555	9	AM423082
34	125	38.7	743	14	CA474136
35	111	34.4	318	14	CB940191
36	100	31.0	380	12	BI883790
37	74	22.9	477	10	AW958112
38	73.5	22.8	584	10	BG078473
39	73.5	22.8	588	12	BM022353
40	73.5	22.8	884	12	BI854371
41	71	22.0	508	12	BJ505938
42	71	22.0	605	12	BU499048
43	69	21.4	631	12	BM18979
44	68.5	21.2	778	14	CB236976
45	68	21.1	803	28	BH443382

DNA ending seq 26 ALIGNMENTS

RESULT 1
LOCUS AA114228 284 bp mRNA linear EST 13-NOV-1996
DEFINITION zn75905.r1 Striatagene NT2 neuronal precursor 937330 Homo sapiens
ACCESSION AA114228
VERSION AA114228.1 GI:1668121
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (Bases 1 to 284)

THIS PAGE BLANK (USPTO)

AUTHORS

Hillier, L., Lennon, G., Becker, M., Bernaldo, M.F., Chiappelli, B., Chissee, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Maritz, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R., and Merritt, M. Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

JOURNAL

Genome Res. 6 (9), 807-828 (1996)

MEDLINE

8889549

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

FEATURES

source

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham.

Location/Qualifiers
1. 284

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:4596579"
/db_xref="taxon:9606"
/clone="IMAGE:564056"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/note="Organ: brain; Vector: pBluescript SK-; Site: 1; EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Uninduced, exponentially growing neuroepithelial cells (Ntera-2/cl.D1). Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"

BASE COUNT 93 a 63 c 56 g 71 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 2,76e-36 Length: 284
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-001-254-26 (1-59) x AAl14228 (1-284)

QY 1 MetAsnLysProlIeThrProSerThyTyrValArgCysLeuAsnValGlyLeuIleArg 20
DB 61 ATGAACAAACCCATCAACATCAATATGCGCTGCTCAATGTTGACTTAATTAGG 120
QY 21 LysLeuSerAspPheIleAspProGlnGlyTyrPylsLysLeuAlaValAlaIleLys 40
DB 121 AAGCTGTAGATTTTATGATTCCTCAAGAGATGAGAGAGTACTGACTATTTAA 180
QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgCysCysSerGlnAsn 59
DB 181 AAACCATCTGCTGATGATAGATCAATCACTTTCACATAGATGCTGTTCCCAAAAC 237

RESULT 2

BF238344

LOCUS

60130461371 NIH_MGC_54 Homo sapiens cDNA clone IMAGE:4132682 5',

DEFINITION

mRNA sequence.

ACCESSION

BF238344

VERSION

BF238344.1

KEYWORDS

EST, Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 576)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished

COMMENT

Contact: Robert Strussberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: InCyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Place: LNCMI032 row: 1 column: 03

High quality sequence start: 6

High quality sequence stop: 536.

FEATURES

source

Location/Qualifiers

1. 576
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4132682"
/tissue_type="from chronic myelogenous leukemia"
/lab_host="RDH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_54"
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech), Site 1: SfiI (ggcgccctcgcc); Site 2: SfiI (ggccatcgcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCTAGAGCCGCGGCGCCGCGCATG-dt(30)-BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 191 a 108 c 132 g 145 t
ORIGIN

Alignment Scores:

Pred. No.: 5.43e-32 Length: 576
Score: 296.00 Matches: 54
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.18% Mismatches: 0
Query Match: 91.64% Indels: 0
DB: 10 Gaps: 0

US-10-001-254-26 (1-59) x BF238344 (1-576)

QY 1 MetAsnLysProlIeThrProSerThyTyrValArgCysLeuAsnValGlyLeuIleArg 20
DB 68 ATGAACAAACCCATCAACATCAATATGCGCTGCTCAATGTTGACTTAATTAGG 127
QY 21 LysLeuSerAspPheIleAspProGlnGlyTyrPylsLysLeuAlaValAlaIleLys 40
DB 128 AAGCTGTAGATTTTATGATTCCTCAAGAGATGAGAGAGTACTGACTATTTAA 187
QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgCys 55
DB 188 AAACCATCTGCTGATGATAGATCAATCACTTTCACATAGATGCTGTTCCCAAAAC 232

RESULT 3

AL699213

LOCUS

540 bp mRNA linear EST 21-MAR-2002

DEFINITION

DKFZp686K18112.1 686 (synonym: hicc3) Homo sapiens cDNA clone

ACCESSION

AL699213

VERSION

AL699213.1

KEYWORDS

EST, Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 23:50:52 ; Search time 51 Seconds

(without alignments)
2544.444 Million cell updates/sec

Title: US-10-001-254-5

Perfect score: 294
Sequence: 1 acatattgctgcgtgcctcaaa.....tgctccagatgctgtccc 294

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Fred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	294	100.0	833	4 US-09-166-350-10	Sequence 10, Appl
2	40.4	13.7	1879	4 US-09-601-091-1	Sequence 1, Appl
3	40.4	13.7	2024	4 US-09-601-091-3	Sequence 3, Appl
4	40.4	13.7	2024	4 US-09-388-385A-51	Sequence 51, Appl
5	40.4	13.7	2024	4 US-09-887-586A-51	Sequence 51, Appl
6	40.4	13.7	2024	4 US-09-895-752-51	Sequence 51, Appl
7	40.4	13.7	2024	4 US-09-903-012B-51	Sequence 51, Appl
8	33.8	11.5	580073	4 US-08-545-528D-1	Sequence 1, Appl
9	33.6	11.4	2393	4 US-08-961-527-90	Sequence 90, Appl
10	32.8	11.2	202001	4 US-09-734-674-3	Sequence 3, Appl
11	32.6	11.1	1728	4 US-09-036-731A-1	Sequence 1, Appl
12	32.4	11.0	98844	4 US-09-791-211-10	Sequence 10, Appl
13	32.2	11.0	4402	4 US-09-484-970B-135	Sequence 135, Appl
14	32	10.9	2274	4 US-09-107-532A-3535	Sequence 3535, Appl
15	31.8	10.8	1557	4 US-09-134-001C-1614	Sequence 1614, Appl
16	31.8	10.8	2507	4 US-09-620-312D-120	Sequence 120, Appl
17	31.8	10.8	1664976	4 US-08-516-421B-1	Sequence 1, Appl
18	31.4	10.7	12720	1 US-08-403-866-11	Sequence 11, Appl
19	31	10.5	2126	2 US-08-545-745-1	Sequence 1, Appl
20	30.8	10.5	1550	4 US-09-427-501-1	Sequence 1, Appl
21	30.8	10.5	15016	4 US-09-601-198-60	Sequence 60, Appl
22	30.8	10.5	319608	4 US-09-539-333D-1	Sequence 1, Appl
23	30.8	10.5	319608	4 US-09-679-409-1	Sequence 1, Appl
24	30.8	10.5	1230025	4 US-09-198-452A-1	Sequence 1, Appl
25	30.6	10.4	1479	4 US-09-328-352-897	Sequence 897, Appl
26	30.4	10.3	1362	3 US-08-929-329-4	Sequence 4, Appl
27	30.2	10.3	2254	2 US-08-635-066-1	Sequence 1, Appl

C 28	30	10.2	195	1 US-08-324-243-20	Sequence 20, Appl
C 29	30	10.2	195	1 US-08-532-390-20	Sequence 20, Appl
C 30	30	10.2	195	1 US-08-717-294-20	Sequence 20, Appl
C 31	30	10.2	195	5 PCT-US95-11511-20	Sequence 20, Appl
C 32	30	10.2	1632	1 US-08-324-243-34	Sequence 34, Appl
C 33	30	10.2	1632	1 US-08-532-390-34	Sequence 34, Appl
C 34	30	10.2	1632	5 US-08-717-294-34	Sequence 34, Appl
C 35	30	10.2	1632	5 PCT-US95-11511-34	Sequence 34, Appl
C 36	30	10.2	2481	1 US-08-324-243-35	Sequence 35, Appl
C 37	30	10.2	2481	1 US-08-532-390-35	Sequence 35, Appl
C 38	30	10.2	2481	3 US-08-717-294-35	Sequence 35, Appl
C 39	30	10.2	2481	5 PCT-US95-11511-35	Sequence 35, Appl
C 40	30	10.2	2571	4 US-09-552-950-4	Sequence 4, Appl
C 41	30	10.2	11770	4 US-08-961-527-172	Sequence 24, Appl
C 42	29.8	10.1	4648	4 US-09-207-914-24	Sequence 24, Appl
C 43	29.8	10.1	4655	4 US-09-207-914-24	Sequence 24, Appl
C 44	29.6	10.1	1356	1 US-08-592-936B-22	Sequence 22, Appl
C 45	29.6	10.1	1356	1 US-08-788-928A-2	Sequence 2, Appl

ALIGNMENTS

DNA of Seq ID#5

RESULT 1	US-09-166-350-10	Application US/09166350A
Sequence 10	6440663	
Patent No.	6440663	
GENERAL INFORMATION:		
APPLICANT:	Scanlan, Matthew	
APPLICANT:	Chen, Yao	
APPLICANT:	Stocker, Elisabeth	
APPLICANT:	Old, Lloyd	
APPLICANT:	Jager, Elke	
APPLICANT:	Knuth, Alex	
TITLE OF INVENTION:	Renal Cancer Associated Antigens and	
TITLE OF INVENTION:	Uses Therefor	
FILE REFERENCE:	L0461/7051	
CURRENT APPLICATION NUMBER:	US/09/166,350A	
CURRENT FILING DATE:	1998-10-05	
EARLIER APPLICATION NUMBER:	US 09/166,350	
EARLIER FILING DATE:	1998-10-05	
NUMBER OF SEQ ID NOS:	35	
SOFTWARE:	FastSeq for Windows Version 3.0	
SEQ ID NO 10		
LENGTH:	833	
TYPE:	DNA	
ORGANISM:	Homo sapiens	
US-09-166-350-10		
Query Match	100.0%; Score 294; DB 4; Length 833;	
Best Local Similarity	100.0%; Pred. No. 3.4e-82;	
Matches 294; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
1	ACATATGCGCGTGGCTCAATGTTGACTAATTTAGGAAGCTGCAGATTTTATGATCCT	60
74	ACATATGCGCGTGGCTCAATGTTGACTAATTTAGGAAGCTGCAGATTTTATGATCCT	133
61	CAGAAGAGTGAAGAAAGTTAGCTTACTTATTAACCATCTGGTATGATGATAC	120
134	CAGAAGAGTGAAGAAAGTTAGCTTACTTATTAACCATCTGGTATGATGATAC	193
121	AATCAGTTTCATTAAGAGATTTGAACTTACTTCAAACTGGAAAAGTCCACTTCT	180
194	AATCAGTTTCATTAAGAGATTTGAACTTACTTCAAACTGGAAAAGTCCACTTCT	253
181	GAATACGTTTACTGATGGGGGACCAAAATGACAGTTGGTATCTTGGATCTTTTG	240
254	GAATACGTTTACTGATGGGGGACCAAAATGACAGTTGGTATCTTGGATCTTTTG	313
241	ATCCAAATGATTTTGTCTCTGCGAGTCTTTGTCTCCAGATGCTGTTCC	294
314	ATCCAAATGATTTTGTCTCTGCGAGTCTTTGTCTCCAGATGCTGTTCC	367

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 18, 2004, 01:16:08 ; Search time 25.5721 Seconds

(without alignments)
1691.512 Million cell updates/sec

Title: US-10-001-254-6

Perfect score: 517

Sequence: 1 TYVACLVNGLIRKLSDFIDP.....LLIONEFAPASLLPDPAVP 98

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Search: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1h
-Q=/cgn2_1/USPTO.spool/US10001254/runat_16012004_152425_19740/app.query.fasta_1.1109
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=ini -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=43 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10001254@cgn2_1_1_95@runat_16012004_152425_19740 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued_Patents_NA.*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	517	100.0	833	4	US-09-166-350-10
2	96.5	18.7	2288	3	US-09-135-232-1
3	96.5	18.7	2288	4	US-09-863-549-1
4	79	15.3	1806	2	US-08-980-060-1
5	79	15.3	1806	3	US-09-307-185-1
6	79	15.3	3459	2	US-08-980-060-3
7	79	15.3	3459	3	US-09-307-185-3
8	79	15.2	479	2	US-08-980-060-14
9	78.5	15.2	479	3	US-09-307-185-14
10	72	13.9	536165	4	US-09-214-808-1
11	69.5	13.4	1416	3	US-09-234-393-26
12	69.5	13.4	1416	4	US-09-360-545-19

13	69.5	13.4	1416	4	US-09-865-171-26	Sequence 26, Appl
14	69.5	13.4	1785	3	US-09-234-393-49	Sequence 49, Appl
15	69.5	13.4	1785	3	US-09-234-393-51	Sequence 51, Appl
16	69.5	13.4	1785	3	US-09-234-393-53	Sequence 53, Appl
17	69.5	13.4	1785	4	US-09-865-171-49	Sequence 49, Appl
18	69.5	13.4	1785	4	US-09-865-171-51	Sequence 51, Appl
19	69.5	13.4	1785	4	US-09-865-171-53	Sequence 53, Appl
20	69.5	13.4	1785	4	US-09-398-395A-49	Sequence 49, Appl
21	69.5	13.4	1785	4	US-09-887-586A-49	Sequence 49, Appl
22	69.5	13.4	1785	4	US-09-895-752-49	Sequence 49, Appl
23	69.5	13.4	1785	4	US-09-903-012B-49	Sequence 49, Appl
24	69.5	13.4	1977	3	US-09-234-393-23	Sequence 23, Appl
25	69.5	13.4	1977	3	US-09-865-171-23	Sequence 23, Appl
26	66.5	12.9	265	4	US-09-313-294A-3660	Sequence 3660, Ap
27	66	12.8	987	4	US-09-252-991A-11251	Sequence 11251, A
28	66	12.8	1140	4	US-09-252-991A-11181	Sequence 11181, A
29	66	12.8	3141	4	US-09-513-783A-10937	Sequence 10937, A
30	64.5	12.5	1623	4	US-09-513-783A-33	Sequence 33, Appl
31	64.5	12.5	1521	4	US-09-134-001C-470	Sequence 470, Appl
32	63.5	12.3	1473	3	US-08-907-740-6	Sequence 6, Appl
33	63.5	12.3	3059	4	US-09-620-312D-213	Sequence 213, App
34	63.5	12.3	3068	4	US-09-620-312D-215	Sequence 215, App
35	63.5	12.3	3071	4	US-09-620-312D-212	Sequence 212, App
36	63.5	12.3	3134	4	US-09-620-312D-211	Sequence 211, App
37	63	12.2	2472	3	US-08-743-168B-37	Sequence 37, Appl
38	63	12.2	2472	3	US-08-743-168B-35	Sequence 35, Appl
39	63	12.2	7881	2	US-08-751-189-1	Sequence 1, Appl
40	63	12.2	7881	2	US-09-060-836-1	Sequence 1, Appl
41	63	12.2	7881	3	US-09-184-445-1	Sequence 1, Appl
42	62.5	12.1	1637	4	US-09-724-623-62	Sequence 62, Appl
43	62.5	12.1	4935	2	US-08-631-097-3	Sequence 3, Appl
44	62.5	12.1	5886	2	US-08-810-712-9	Sequence 9, Appl
45	62.5	12.1	6027	2	US-08-968-542C-1	Sequence 1, Appl

ALIGNMENTS

DNA encoding Seq ID #6

RESULT 1
US-09-166-350-10
Sequence 10, Application US/09166350A
Patent No. 6440663
GENERAL INFORMATION:
APPLICANT: Scamian, Matthew
APPLICANT: Chen, Yao
APPLICANT: Stockert, Elisabeth
APPLICANT: Old, Lloyd
APPLICANT: Jager, Elke
APPLICANT: Knuth, Alex
TITLE OF INVENTION: Renal Cancer Associated Antigens and
FILE REFERENCE: 10461/7051
CURRENT APPLICATION NUMBER: US/09/166,350A
CURRENT FILING DATE: 1998-10-05
EARLIER APPLICATION NUMBER: US 09/166,350
EARLIER FILING DATE: 1998-10-05
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 833
TYPE: DNA
ORGANISM: Homo sapiens
US-09-166-350-10
Alignment Scores:
Pred. No.: 1,176-67
Score: 517.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Length: 833
Matches: 98
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0
US-10-001-254-6 (1-98) x US-09-166-350-10 (1-833)

THIS PAGE BLANK (USPTO)

QY 1 ThrTyrValArgCysLeuSerValGlyLeuLeuArgLysLeuSerAspPheIleAspPro 20
Db 74 ACATATGTCGGCTGCTCATATGTGACAAATAGAGAGCTGCACATTTTATTGATCCT 133
QY 21 GlnGlnGlyTyrPheLysLeuAlaValAlaIleLysLysProSerGlyAspAspArgTyr 40
Db 134 CAGAGAGATGAGAGAGATTAGCTGATGATTTAAAAACCATCTCTGATGATGATGATC 193
QY 41 AsnGlnPheHisIleArgArgPheGlnAlaLeuLeuGlnThrGlyLysSerProThrSer 60
Db 194 AATCAGCTTCACATTAAGAGATTGAGACATTACTTCAACTGAGAAAAAGTCCACTTCT 253
QY 61 GlnLeuLeuPheSerTyrPheGlyThrThrAsnCysThrValGlyAspLeuValAspLeu 80
Db 254 GAATTAAGCTGTGACGCGGACCAAAATGACAGATTGATCTTGTGATCTTTTGG 313
QY 81 IleGlnAsnGlnPhePheAlaProAlaSerLeuLeuProAspAlaValPro 98
Db 314 ATCCAAATGAATTTTCTCTCTGCGAGCTTTTCTCCAGATCTCTTCC 367

RESULT 2

US-09-135-232-1
Sequence 1, Application US/09135232
Patent No. 6262228

GENERAL INFORMATION:

APPLICANT: Cao, Zhaoan
TITLE OF INVENTION: IRAK3 Polypeptides, Polynucleotides and Methods
FILE REFERENCE: T98-019
CURRENT APPLICATION NUMBER: US/09/135,232
CURRENT FILING DATE: 1998-08-17
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 2288
TYPE: DNA
ORGANISM: human
FEATURE:
NAME/KEY: CDS
LOCATION: (64)..(1851)
US-09-135-232-1

Alignment Scores:

Pred. No.: 0.000283 Length: 2288
Score: 96.50 Matches: 23
Percent Similarity: 47.95% Conservative: 12
Best Local Similarity: 31.51% Mismatches: 31
Query Match: 18.67% Indels: 7
DB: 3 Gaps: 1

US-10-001-254-6 (1-98) x US-09-135-232-1 (1-2288)

QY 8 ValGlyLeuIleArgLysLeuSerAspPheIleAspProGlnGlyTyrLysLysLeu 27
Db 136 CTCGGAGACTCTGCGCTGTTCTGACACCTGCGAGCGCGCTGGCGCGGCTG 195
QY 28 AlaValAlaIleLysLysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgArg 47
Db 196 GCAGAGAGACTTCAAGC-----AGCTGGCTGATGATTTGCTCAT 234
QY 48 PheGlnAlaLeuLeuGlnThrGlyLysSerProThrSerGlnLeuLeuPheAspTyrGly 67
Db 235 ATTGAAAAGTATGTAGACCAAGGTAAAGTGAACAAGAGATTACTTTGGTCTGGCA 294
QY 68 ThrThrAsnCysThrValGlyAspLeuValAspLeuLeu 80
Db 295 CAGAAAACAGAGACCATCGGTGACCTTTTACAGGTCTC 333

RESULT 3

US-09-863-549-1
Sequence 1, Application US/09863549
Patent No. 6576444
GENERAL INFORMATION:
APPLICANT: Cao, Zhaoan

TITLE OF INVENTION: IRAK3 Polypeptides, Polynucleotides and Methods
FILE REFERENCE: T98-019
CURRENT APPLICATION NUMBER: US/09/863,549
CURRENT FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: 09/135,232
PRIORITY FILING DATE: 1998-08-17
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 2288
TYPE: DNA
ORGANISM: human
FEATURE:
NAME/KEY: CDS
LOCATION: (64)..(1851)
US-09-863-549-1

Alignment Scores:

Pred. No.: 0.000283 Length: 2288
Score: 96.50 Matches: 23
Percent Similarity: 47.95% Conservative: 12
Best Local Similarity: 31.51% Mismatches: 31
Query Match: 18.67% Indels: 7
DB: 4 Gaps: 1

US-10-001-254-6 (1-98) x US-09-863-549-1 (1-2288)

QY 8 ValGlyLeuIleArgLysLeuSerAspPheIleAspProGlnGlyTyrLysLysLeu 27
Db 136 CTCGGAGACTCTGCGCTGTTCTGACACCTGCGAGCGCGCTGGCGCGGCTG 195
QY 28 AlaValAlaIleLysLysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgArg 47
Db 196 GCAGAGAGACTTCAAGC-----AGCTGGCTGATGATTTGCTCAT 234
QY 48 PheGlnAlaLeuLeuGlnThrGlyLysSerProThrSerGlnLeuLeuPheAspTyrGly 67
Db 235 ATTGAAAAGTATGTAGACCAAGGTAAAGTGAACAAGAGATTACTTTGGTCTGGCA 294
QY 68 ThrThrAsnCysThrValGlyAspLeuValAspLeuLeu 80
Db 295 CAGAAAACAGAGACCATCGGTGACCTTTTACAGGTCTC 333

RESULT 4

US-08-980-060-1
Sequence 1, Application US/08980060
Patent No. 5965421

GENERAL INFORMATION:

APPLICANT: NI, JIAN
APPLICANT: PENG, PING
APPLICANT: MUZIO, MARTA
APPLICANT: DIXIT, VISHVA M.
TITLE OF INVENTION: HUMAN IRAK-2
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P. L. L. C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON
STATE: D. C.
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,060
FILING DATE: Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 18, 2004, 01:16:08 ; Search time 25.5721 Seconds
(without alignments)
1691.512 Million cell updates/sec

Title: US-10-001-254-6
Perfect score: 517
Sequence: 1 TYVRCINWGHIRKLSDFIDP.....LLIQNEFPAPASLLPDPAV 98

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 569978 seqs, 220691566 residues
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFTO.spool/US10001254/rnat_16012004_152425_19740/app_query.fasta_1.1109
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=rml -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsun62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=opt -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=tbl -NOM=ext -HEAPSIZE=500 -MTLEN=0 -MAXLEN=2000000000
-USR=US10001254.@CGN 1.1.95 @rnat_16012004_152425_19740 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	517	100.0	833	4	US-09-166-350-10
2	96.5	18.7	2288	3	US-09-135-232-1
3	96.5	18.7	2288	4	US-09-863-549-1
4	79	15.3	1806	2	US-08-980-060-1
5	79	15.3	1806	3	US-09-307-185-1
6	79	15.3	1806	4	US-09-307-185-1
7	79	15.3	3459	2	US-08-980-060-3
8	79	15.3	3459	3	US-09-307-185-3
9	78.5	15.2	479	2	US-08-980-060-14
10	78.5	15.2	479	3	US-09-307-185-14
11	69.5	13.4	1416	3	US-09-234-393-26
12	69.5	13.4	1416	4	US-09-360-545-19

13	69.5	13.4	1416	4	US-09-865-171-26	Sequence 26, App1
14	69.5	13.4	1785	3	US-09-234-393-49	Sequence 49, App1
15	69.5	13.4	1785	3	US-09-234-393-51	Sequence 51, App1
16	69.5	13.4	1785	3	US-09-234-393-53	Sequence 53, App1
17	69.5	13.4	1785	4	US-09-865-171-49	Sequence 49, App1
18	69.5	13.4	1785	4	US-09-865-171-51	Sequence 51, App1
19	69.5	13.4	1785	4	US-09-865-171-53	Sequence 53, App1
20	69.5	13.4	1785	4	US-09-398-395A-49	Sequence 49, App1
21	69.5	13.4	1785	4	US-09-887-886A-49	Sequence 49, App1
22	69.5	13.4	1785	4	US-09-895-752-49	Sequence 49, App1
23	69.5	13.4	1785	4	US-09-903-012B-49	Sequence 49, App1
24	69.5	13.4	1977	3	US-09-234-393-23	Sequence 23, App1
25	69.5	13.4	1977	4	US-09-865-171-23	Sequence 23, App1
26	66.5	12.9	265	4	US-09-313-294A-3660	Sequence 3660, App1
27	66	12.8	987	4	US-09-252-991A-11251	Sequence 11251, App1
28	66	12.8	1140	4	US-09-252-991A-11181	Sequence 11181, App1
29	66	12.8	1314	4	US-09-252-991A-10937	Sequence 10937, App1
30	65	12.6	1633	4	US-09-513-783A-33	Sequence 33, App1
31	64.5	12.5	1521	4	US-09-134-001C-470	Sequence 470, App1
32	63.5	12.3	1473	3	US-08-907-740-6	Sequence 6, App1
33	63.5	12.3	3059	4	US-09-620-312D-213	Sequence 213, App1
34	63.5	12.3	3068	4	US-09-620-312D-214	Sequence 214, App1
35	63.5	12.3	3071	4	US-09-620-312D-215	Sequence 215, App1
36	63.5	12.3	3134	4	US-09-620-312B-212	Sequence 212, App1
37	63	12.2	2472	3	US-08-743-168B-35	Sequence 35, App1
38	63	12.2	2472	3	US-08-743-168B-37	Sequence 37, App1
39	63	12.2	7881	2	US-08-751-189-1	Sequence 1, App1
40	63	12.2	7881	2	US-09-060-836-1	Sequence 1, App1
41	63	12.2	7881	3	US-09-184-445-1	Sequence 1, App1
42	62.5	12.1	1637	4	US-09-724-623-62	Sequence 62, App1
43	62.5	12.1	4935	2	US-08-631-097-3	Sequence 3, App1
44	62.5	12.1	5886	3	US-08-810-712-9	Sequence 9, App1
45	62.5	12.1	6027	2	US-08-968-542C-1	Sequence 1, App1

ALIGNMENTS

Diffentially Sea ID #6

RESULT 1
US-09-166-350-10
Sequence 10, Application US/09166350A
Patent No. 6440663
GENERAL INFORMATION:
APPLICANT: Scanlan, Matthew
APPLICANT: Stockert, Elisabeth
APPLICANT: Old, Lloyd
APPLICANT: Jager, Elke
APPLICANT: Knuth, Alex
TITLE OF INVENTION: Renal Cancer Associated Antigens and
TITLE OF INVENTION: Uses Therefor
FILE REFERENCE: L0461/7051
CURRENT APPLICATION NUMBER: US/09/166,350A
CURRENT FILING DATE: 1998-10-05
EARLIER APPLICATION NUMBER: US 09/166,350
EARLIER FILING DATE: 1998-10-05
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 833
TYPE: DNA
ORGANISM: Homo sapiens
US-09-166-350-10
Alignment Scores:
Pred. No.: 1.17e-67
Score: 517.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 4
Gaps: 0
US-10-001-254-6 (1-98) x US-09-166-350-10 (1-833)

```
QY      1 ThrTYValArgCysLeuValGlyLeuLeuArgGlyLeuSerAspPheIleAspPro 20
      |||
      74 AARATATGCGCTGCTCATATGTGACATTAAGAACTGTGACATTTATGATCTT 133
QY      21 GINGUGLYTTPlyLysLeuValAlaAlaIleLysLysProSerGlyAspAspArgTYR 40
      |||
      134 CAAGAAGATGAGAAAGATTAGCTGTAGCTATTAATAAACCAATCTGTATGATGATAC 193
QY      41 AANGINPheHsiIleArgArgPheGlyAlaLeuLeuGlnThrGlyLysSerProThrSer 60
      |||
      194 AATCAGTTTCACATTAAGAGATTGAAGCATTACTTCAAACTGGAATAAGTCCACTTCT 253
QY      61 GluLeuLeuPheAspTyrGlyThrThrAsnCysThrValGlyAspLeuValAspLeuLeu 80
      |||
      254 GAATTCGTGTTGACGTGGGACCAACAATGACAGCTTGTCATCTTGTGATCTTTTG 313
QY      81 IlegInAsnGluPhePheAlaProAlaSerLeuLeuProAspAlaValPro 98
      |||
      314 ATCCAAATGAATTTTGTGCTCTCGAGCTTTTCTCCAGATGCTGTTCCC 367

RESULT 2
US-09-135-232-1
; Sequence 1, Application US/09135232
; Patent No. 6262228
; GENERAL INFORMATION:
; APPLICANT: Cao, Zhaoan
; TITLE OF INVENTION: IRAK3 Polypeptides, Polynucleotides and Methods
; FILE REFERENCE: T98-019
; CURRENT APPLICATION NUMBER: US/09/135,232
; CURRENT FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2288
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (64)..(1851)
US-09-135-232-1

Alignment Scores:
Pred. No.: 0.000283 Length: 2288
Score: 96.50 Matches: 23
Percent Similarity: 47.95% Conservative: 12
Best Local Similarity: 31.51% Mismatches: 31
Query Match: 18.67% Indels: 7
DB: 3 Gaps: 1

US-10-001-254-6 (1-98) x US-09-135-232-1 (1-2288)
QY      8 ValGlyLeuIleArgLysLeuSerAspPheIleAspProGlnGlnGlyTTPlyLysLeu 27
      |||
      136 CTCGGAGAGCTCGCTGTCGACAGCTGCGCGCGCTGGCGCGCTGG 195
QY      28 AlaValAlaIleLysLysProSerGlyAspAspArgTYRAsnGlnPheHsiIleArgArg 47
      |||
      196 GCAGAGAGACTTCAAGC-----AGCTGGCTGAGTGTGTCAT 234
QY      48 PheGluAlaLeuLeuGlnThrGlyLysSerProThrSerGluLeuLeuPheAspTyrGly 67
      |||
      235 ATTGAAGAGTATGTAACCAAGGTAAAGTGAACAAGAAATTAATCTTGTGCTGGCA 294
QY      68 ThrThrAsnCysThrValGlyAspLeuValAspLeuLeu 80
      |||
      295 CAGAAAAACAAAGACCATCGGTGACCTTTTACAGGTCTC 333
DB

RESULT 3
US-09-863-549-1
; Sequence 1, Application US/09863549
; Patent No. 6576444
; GENERAL INFORMATION:
; APPLICANT: Cao, Zhaoan
```

```
; TITLE OF INVENTION: IRAK3 Polypeptides, Polynucleotides and Methods
; FILE REFERENCE: T98-019
; CURRENT APPLICATION NUMBER: US/09/863,549
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 09/135,232
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2288
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (64)..(1851)
US-09-863-549-1

Alignment Scores:
Pred. No.: 0.000283 Length: 2288
Score: 96.50 Matches: 23
Percent Similarity: 47.95% Conservative: 12
Best Local Similarity: 31.51% Mismatches: 31
Query Match: 18.67% Indels: 7
DB: 4 Gaps: 1

US-10-001-254-6 (1-98) x US-09-863-549-1 (1-2288)
QY      8 ValGlyLeuIleArgLysLeuSerAspPheIleAspProGlnGlnGlyTTPlyLysLeu 27
      |||
      136 CTCGGAGAGCTCGCTGTCGACAGCTGCGCGCGCTGGCGCGCTGG 195
QY      28 AlaValAlaIleLysLysProSerGlyAspAspArgTYRAsnGlnPheHsiIleArgArg 47
      |||
      196 GCAGAGAGACTTCAAGC-----AGCTGGCTGAGTGTGTCAT 234
QY      48 PheGluAlaLeuLeuGlnThrGlyLysSerProThrSerGluLeuLeuPheAspTyrGly 67
      |||
      235 ATTGAAGAGTATGTAACCAAGGTAAAGTGAACAAGAAATTAATCTTGTGCTGGCA 294
QY      68 ThrThrAsnCysThrValGlyAspLeuValAspLeuLeu 80
      |||
      295 CAGAAAAACAAAGACCATCGGTGACCTTTTACAGGTCTC 333
DB

RESULT 4
US-08-980-060-1
; Sequence 1, Application US/08980060
; Patent No. 5965421
; GENERAL INFORMATION:
; APPLICANT: NI, JIAN
; APPLICANT: FENG, PING
; APPLICANT: MUZIO, MARTA
; APPLICANT: DIXIT, VISHVA M.
; TITLE OF INVENTION: HUMAN IRAK-2
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/980,060
; FILING DATE: Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
```

```

1  REFERENCE/DOCKET NUMBER: 1488.1010000/EKS/ALJK
2  TELECOMMUNICATION INFORMATION:
3  TELEPHONE: (202) 371-2500
4  TELEFAX: (202) 371-2540
5  INFORMATION FOR SEQ ID NO: 1:
6  SEQUENCE CHARACTERISTICS:
7  LENGTH: 1806 base pairs
8  TYPE: nucleic acid
9  STRANDEDNESS: double
10 TOPOLOGY: linear
11 MOLECULE TYPE: DNA (genomic)
12 FEATURE:
13 NAME/KEY: CDS
14 LOCATION: 34..1803
15 US-08-980-060-1
16
17 Alignment Scores:
18 Pred. No.: 0.0822 Length: 1806
19 Score: 79.00 Matches: 22
20 Percent Similarity: 49.28% Conservative: 12
21 Best Local Similarity: 31.88% Mismatches: 19
22 Query Match: 15.28% Indels: 16
23 DB: 2 Gaps: 2
24
25 US-10-001-254-6 (1-98) x US-08-980-060-1 (1-1806)
26
27 QY 45 IleaArgArgPheGluAlaLeuLeuGlnThr--GlyValSerProThrSerGluLeuLeu 63
28      CCGGCGGAGAGTCAACATCCATCATGAGCGGGTGCAGAGGTGTGAGCATCAGCGGGAGCTGCTG 21
29      ::::|||||:::|||||
30 Db 157 TGGGCGGAGAGTCAACATCCATCATGAGCGGGTGCAGAGGTGTGAGCATCAGCGGGAGCTGCTG 21
31      ::::|||||:::|||||
32 QY 64 PheAspIrrpGlyThrThrValencySThrValIGlyAspLeuValAspLeuLeuLeuGlnAsn 83
33      TGGGCGGAGAGTCAACATCCATCATGAGCGGGTGCAGAGGTGTGAGCATCAGCGGGAGCTGCTG 27
34      ::::|||||:::|||||
35 Db 217 TGGGCGGAGAGTCAACATCCATCATGAGCGGGTGCAGAGGTGTGAGCATCAGCGGGAGCTGCTG 27
36      ::::|||||:::|||||
37 QY 84 GluPhePheAlaProAlaSerLeuLeuLeu----- 93
38      |||:::|||||
39 Db 277 GAGCTCTACCGGGCGGTGCGCCAGATCATCTGTGAAGTGAACCGGCTCTGTAATCAAGGTGT 33
40      |||:::|||||
41 QY 94 -----ProAspAlaVal 97
42      |||:::|||||
43 Db 337 CCCATTCACGCTTCCCTGACTCTGTG 363
44
45 RESULT 5
46 US-09-307-185-1
47 Sequence 1, Application US/09307185
48 Patent No. 6222019
49 GENERAL INFORMATION:
50 APPLICANT: NI, JIAN
51 APPLICANT: PENG, PING
52 APPLICANT: MUZIO, MARTA
53 APPLICANT: DIXIT, VISHVA M.
54 TITLE OF INVENTION: HUMAN IRAK-2
55 NUMBER OF SEQUENCES: 14
56 CORRESPONDENCE ADDRESS:
57 ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
58 STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
59 CITY: WASHINGTON
60 STATE: D.C.
61 COUNTRY: USA
62 ZIP: 20005-3934
63 COMPUTER READABLE FORM:
64 MEDIUM TYPE: Floppy disk
65 COMPUTER: IBM PC compatible
66 OPERATING SYSTEM: PC-DOS/MS-DOS
67 SOFTWARE: Patent Release #1.0, Version #1.30
68 CURRENT APPLICATION DATA:
69 APPLICATION NUMBER: US/09307,185
70 FILING DATE:
71 CLASSIFICATION:
72 PRIOR APPLICATION DATA:
73 APPLICATION NUMBER: US/08/980,060
74 FILING DATE:
75 ATTORNEY/AGENT INFORMATION:

```

```

1 NAME: STEFFE, ERIC K.
2 REGISTRATION NUMBER: 36,688
3 REFERENCE/DOCKET NUMBER: 1488.1010000/EKS/AJK
4 TELECOMMUNICATION INFORMATION:
5 TELEPHONE: (202) 371-2600
6 TELEFAX: (202) 371-2540
7 INFORMATION FOR SEQ ID NO: 1:
8 SEQUENCE CHARACTERISTICS:
9 LENGTH: 1806 base pairs
10 TYPE: nucleic acid
11 STRANDEDNESS: double
12 TOPOLOGY: linear
13 MOLECULE TYPE: DNA (genomic)
14 FEATURE:
15 NAME/KEY: CDS
16 LOCATION: 34..1803
17 US-09-307-185-1
18
19 Alignment Scores:
20 Pred. No.: 0.0822 Length: 1806
21 Score: 79.00 Matches: 12
22 Percent Similarity: 49.28% Conservative: 22
23 Best Local Similarity: 31.88% Mismatches: 19
24 Query Match: 15.28% Indels: 16
25 DB: 3 Gaps: 2
26
27 US-10-001-254-6 (1-98) x US-09-307-185-1 (1-1806)
28
29 Oy 45 ILAARGARphEglnAlaLeuengInThr---GlyLysSerProThrSergLLeuLeu 63
30 Db 157 CTGCGAAGATCAAGTCATGAGCGCGGTCGAGCATCACCAGGAGCTGCTG 216
31 Oy 64 PheAspTrpGlyThrThraSncYsThrValGlyAspLeuValAspLeuLeuIleGlnAsn 83
32 Db 217 TGGTGGGAGGCGATGCGGCGAGGCCACCGTCGACCACTTGAGACTCTGTGCGGCTG 276
33 Oy 84 GluPheAlaIraProAlaSerLeuLeu----- 93
34 Db 277 GAGCTTACCGGGCTGCCCAATCATCTGAACTGGAACCGGCTCTGAATCAGGTGT 336
35 Oy 94 -----ProAspAlaVal 97
36 Db 337 CCCATTCCAGCTTCCCTGACTGTGTG 363
37
38 RESULT 6
39 US-08-980-060-3
40 Sequence 3, Application US/08980060
41 Patent No. 5965421
42 GENERAL INFORMATION:
43 APPLICANT: NI, JIAN
44 APPLICANT: FENG, PING
45 APPLICANT: MUZIO, MARTA
46 APPLICANT: DIXIT, VISHVA M.
47 TITLE OF INVENTION: HUMAN IRAK-2
48 NUMBER OF SEQUENCES: 14
49 CORRESPONDENCE ADDRESSES:
50 ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
51 STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
52 CITY: WASHINGTON
53 STATE: D.C.
54 COUNTRY: USA
55 ZIP: 20005-3934
56 COMPUTER READABLE FORM:
57 MEDIUM TYPE: Floppy disk
58 COMPUTER: IBM PC compatible
59 OPERATING SYSTEM: PC-DOS/MS-DOS
60 SOFTWARE: PatentIn Release #1.0, Version #1.30
61 CURRENT APPLICATION DATA:
62 APPLICATION NUMBER: US/08/980,060
63 FILING DATE: Herewith
64 CLASSIFICATION: A35
65 ATTORNEY/AGENT INFORMATION:
66 NAME: STEFFE, ERIC K.

```

REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1010000/EKS/AJK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3459 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 34..1908
US-08-980-060-3

Alignment Scores:
Pred. No.: 0.211 Length: 3459
Score: 79.00 Matches: 22
Percent Similarity: 49.28% Conservative: 12
Best Local Similarity: 31.88% Mismatches: 19
Query Match: 15.28% Indels: 16
Gaps: 2

US-10-001-254-6 (1-98) x US-08-980-060-3 (1-3459)

QY 45 ILAARGAphghegluaLauleuGlnThr---GlyVSSerProThrsrGluLeuLeu 63
DB 157 CTGCGGAAAGATCAAGTCATGAGCGGGTGTGAGCATCAGCGGAGAGCTGCTG 216
QY 64 PhAspTrpGlyThrThrsnCyThrValGlyAspleuValAspleuLeuIleGlnAsn 83
DB 217 TGGTGTGGGGCATGCGGCGAGCCACCCCTCCAGAACTTGTGACCTCTGTGCGCGCTG 276
QY 84 GluphePhaLaProLaSerLeuLeuLeu----- 93
DB 277 GAGCTTACCGGGCTGCCAGATCATCTGAACTGMAACCGGCTCCTGAATCAGGTGT 336
QY 94 -----ProAspAlaVal 97
DB 337 CCCATTCCAGCCTTCCCTGACTGTGTG 363

RESULT 7
US-09-307-185-3
Sequence 3, Application US/09307185
Patent No. 6222019
GENERAL INFORMATION:
APPLICANT: NI, JIAN
APPLICANT: FENG, PING
APPLICANT: MUZIO, MARTA
APPLICANT: DIXIT, VISHVA M.
TITLE OF INVENTION: HUMAN IRK-2
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/307,185
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/980,060
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1010000/EKS/AJK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3459 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 34..1908
US-09-307-185-3

Alignment Scores:
Pred. No.: 0.211 Length: 3459
Score: 79.00 Matches: 22
Percent Similarity: 49.28% Conservative: 12
Best Local Similarity: 31.88% Mismatches: 19
Query Match: 15.28% Indels: 16
Gaps: 2

US-10-001-254-6 (1-98) x US-09-307-185-3 (1-3459)

QY 45 ILAARGAphghegluaLauleuGlnThr---GlyVSSerProThrsrGluLeuLeu 63
DB 157 CTGCGGAAAGATCAAGTCATGAGCGGGTGTGAGCATCAGCGGAGAGCTGCTG 216
QY 64 PhAspTrpGlyThrThrsnCyThrValGlyAspleuValAspleuLeuIleGlnAsn 83
DB 217 TGGTGTGGGGCATGCGGCGAGCCACCCCTCCAGAACTTGTGACCTCTGTGCGCGCTG 276
QY 84 GluphePhaLaProLaSerLeuLeuLeu----- 93
DB 277 GAGCTTACCGGGCTGCCAGATCATCTGAACTGMAACCGGCTCCTGAATCAGGTGT 336
QY 94 -----ProAspAlaVal 97
DB 337 CCCATTCCAGCCTTCCCTGACTGTGTG 363

RESULT 8
US-08-980-060-14/c
Sequence 14, Application US/08980060
Patent No. 5965421
GENERAL INFORMATION:
APPLICANT: NI, JIAN
APPLICANT: FENG, PING
APPLICANT: MUZIO, MARTA
APPLICANT: DIXIT, VISHVA M.
TITLE OF INVENTION: HUMAN IRK-2
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,060
FILING DATE: Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

	TELEFAX: (202) 371-2540		
	INFORMATION FOR SEQ ID NO: 14:		
	SEQUENCE CHARACTERISTICS:		
	LENGTH: 479 base pairs		
	TYPE: nucleic acid		
	STRANDEDNESS: double		
	TOPOLOGY: linear		
	MOLECULE TYPE: cDNA		
	US-09-307-185-14		
	Alignment Scores:		
Pred. No.:	0.0143	Length:	479
Score:	78.50	Matches:	19
Percent Similarity:	60.00%	Conservative:	11
Best Local Similarity:	38.00%	Mismatches:	19
Query Match:	15.18%	Indels:	1
DB:	3	Gaps:	1
US-10-001-254-6 (1-98) x US-09-307-185-14 (1-479)			
OY	45	11leAgtArpPhegiAlalaLeuLglnThr---	GlyLySerProThrSerGluLeu 63
	:::::	:	:
Db	244	CTGGGAAGATCAAGTCATGAGCGGTGCAGGGTGTGACATCAGCGGAGACTGCTG	185
OY	64	PheAsPrIcglYthrThrAsnCySrmVal(GlyAspleuValaspleuleuIIeGIAsn	83
	:::	:	:
Db	184	TGGTGtGGGCGATGCGGACGCCACCCTCCAGCAACTGTGTGACCTCTGTGCCGCTG	125
OY	84	GluPheheaIaProAlaSerLeuLeu 93	
		:	:
Db	124	GAGCTTAACGGGCTGCCAGATCATCTTG	95
RESULT 10			
US-09-214-808-1			
; Sequence 1, Application US/09214808A			
; Patent No. 6475793			
; GENERAL INFORMATION:			
; APPLICANT: Rosenthal, Andre			
; APPLICANT: Freiberg, Christoph			
; APPLICANT: Perret, Xavier Philippe			
; APPLICANT: Broughton, William John			
; TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic			
; Patent No. 6475793			
; TITLE OF INVENTION: Plasmid			
; FILE REFERENCE: CARP0068			
; CURRENT APPLICATION NUMBER: US/09/214, 808A			
; CURRENT FILING DATE: 1999-06-22			
; PRIOR APPLICATION NUMBER: PCT/IB97/00950			
; PRIOR FILING DATE: 1997-07-10			
; NUMBER OF SEQ ID NOS: 1			
; SOFTWARE: Patentin Ver. 2.1			
; SEQ ID NO 1			
LENGTH: 536165			
TYPE: DNA			
ORGANISM: Rhizobium			
US-09-214-808-1			
Alignment Scores:			
Pred. No.:	3.46e+03	Length:	536165
Score:	72.00	Matches:	28
Percent Similarity:	47.06%	Conservative:	12
Best Local Similarity:	32.94%	Mismatches:	39
Query Match:	13.93%	Indels:	6
DB:	4	Gaps:	5
US-10-001-254-6 (1-98) x US-09-214-808-1 (1-536165)			
OY	18	11leAsPProGIngluGIYTP--TyRbLYleuAlaValAl1leLYslySProSer---	35
		:	:
Db	97239	ATTGACGATCGGAGACTTGCGGACAGAAAATCATCGTTGCTGCTCAAGTCCGTGCG	97299
OY	36	--GlyAspAsPrAr--TyRbngInPheHs11leAgtArpPhegiAlalaLeuLgln	53
		:	:

```

Db      97299 ACCGGGATGACATCTGACTGCTGCTCCGACAGATCCACAGATATCATCTTGTGTAG 97358
QY      54 ThrGlyLysSerProthrSerGluLeuLeuPheAspTrrGlyThrThrAsnCysThrVal 73
Db      97359 ATTGAGCCTTCGGCTCACTTACCTTGTGATGG---TCTACTACTGCGCTCTCA 97415
QY      74 ---GlyAspLeuValAspLeuLeuIleGlnAsnGluPhePheAlaProAlaSerLeuLeu 92
Db      97416 TGAAGGACCTTACAGCGCGCTCTTCCATCGCTTCTTATCCGCTCTTCAAAATATG 97475
QY      93 LeuProAspAlaVal 97
Db      97476 GTGCAATCATTTGTT 97490

RESULT 11
US-09-234-393-26
/ Sequence 26, Application US/09234393A
/ Patent No. 6265639
/ GENERAL INFORMATION:
/ APPLICANT: Croteau, Rodney B
/ APPLICANT: Croteau, John E
/ APPLICANT: Bohlman, Jorg
/ APPLICANT: Jetter, Reinhard
/ APPLICANT: Steele, Christopher L
/ TITLE OF INVENTION: SEQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)
/ FILE REFERENCE: WSUR113345
/ CURRENT APPLICATION NUMBER: US/09/234,393A
/ EARLIER FILING DATE: 1999-01-20
/ EARLIER APPLICATION NUMBER: 60/072,204
/ NUMBER OF SEQ ID NOS: 55
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 26
/ LENGTH: 1416
/ TYPE: DNA
/ ORGANISM: Abies grandis
US-09-234-393-26

Alignment Scores:
Pred. No.: 1.51 Length: 1416
Score: 69.50 Matches: 41
Percent Similarity: 40.91% Conservative: 13
Best Local Similarity: 31.06% Mismatches: 37
Query Match: 13.44% Indels: 41
DB: 3 Gaps: 10

US-10-001-254-6 (1-98) x US-09-234-393-26 (1-1416)
QY      3 ValArgCysLeuAsnValGlyLeuIleArgLysLeuSerAspPheIleAsp---ProGln 21
Db      501 GTGAGACGATGGAGATGTTTGTGTGTAAGGCGCTCCAGACTTCATGAAATTCATTC 560
QY      22 GluGlyTrpLysLys-----LeuAlaValAlaIleLysLysProSergly 36
Db      561 GAGTTCGTGTTAAAGACATCTAATGATGATGCTGAAGCTGT---AAAGCGAAGG 617
QY      37 AspAsp-----ArgTyrAsnGlnPheHisIleArgArgPheGluAlaLeu 51
Db      618 CAAGATATGCGCGCTACATAGAAAATGATGAGAG---CGATACCTTGAAGCTTAT 674
QY      52 LeuGln-----ThrGlyLysSerProthrSerGluLeuLeuPheAsp 65
Db      675 CTGCAAGATCGCGAATGATAGCCATGCGACATGCTCCACCTTTGATGAGTACTTGAAT 734
QY      66 TrpGlyThrThrAsn-----CysThrValGlyAspLeuValAspLeuLeuIleGln 82
Db      735 AATGGACACCAAAACACTGGGATGTGTATG---AATTGATCCGCTTGTGTAAG 791
QY      83 AsnGlu-----PhePheAlaPro----- 88
Db      792 GGTGAACATTTACCAATCGAATTCGTGAGCAAAATATTCTTGCCCTCCAGTTCCACAT 851

```

```

QY      89 -----AlaSerLeuLeuProAspAla 96
Db      852 CTCATTGAATTGGCTTCAGGCTCGTCGATGACGG 887

RESULT 12
US-09-360-545-19
/ Sequence 19, Application US/09360545
/ Patent No. 6429014
/ GENERAL INFORMATION:
/ APPLICANT: Croteau, Rodney B
/ APPLICANT: Bohlman, Jorg
/ APPLICANT: Steele, Christopher L
/ APPLICANT: Phillips, Michael A
/ TITLE OF INVENTION: MONOTERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)
/ FILE REFERENCE: WSUR13885
/ CURRENT APPLICATION NUMBER: US/09/360,545
/ EARLIER FILING DATE: 1999-07-26
/ EARLIER APPLICATION NUMBER: 60/052,249
/ EARLIER FILING DATE: 1997-11-07
/ EARLIER APPLICATION NUMBER: PCT/US98/14528
/ NUMBER OF SEQ ID NOS: 107
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 19
/ LENGTH: 1416
/ TYPE: DNA
/ ORGANISM: Abies grandis
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (3)..(1199)
/ OTHER INFORMATION: Clone AGS.9
US-09-360-545-19

Alignment Scores:
Pred. No.: 1.51 Length: 1416
Score: 69.50 Matches: 41
Percent Similarity: 40.91% Conservative: 13
Best Local Similarity: 31.06% Mismatches: 37
Query Match: 13.44% Indels: 41
DB: 4 Gaps: 10

US-10-001-254-6 (1-98) x US-09-360-545-19 (1-1416)
QY      3 ValArgCysLeuAsnValGlyLeuIleArgLysLeuSerAspPheIleAsp---ProGln 21
Db      501 GTGAGACGATGGAGATGTTTGTGTGTAAGGCGCTCCAGACTTCATGAAATTCATTC 560
QY      22 GluGlyTrpLysLys-----LeuAlaValAlaIleLysLysProSergly 36
Db      561 GAGTTCGTGTTAAAGACATCTAATGATGATGCTGAAGCTGT---AAAGCGAAGG 617
QY      37 AspAsp-----ArgTyrAsnGlnPheHisIleArgArgPheGluAlaLeu 51
Db      618 CAAGATATGCGCGCTACATAGAAAATGATGAGAG---CGATACCTTGAAGCTTAT 674
QY      52 LeuGln-----ThrGlyLysSerProthrSerGluLeuLeuPheAsp 65
Db      675 CTGCAAGATCGCGAATGATAGCCATGCGACATGCTCCACCTTTGATGAGTACTTGAAT 734
QY      66 TrpGlyThrThrAsn-----CysThrValGlyAspLeuValAspLeuLeuIleGln 82
Db      735 AATGGACACCAAAACACTGGGATGTGTATG---AATTGATCCGCTTGTGTAAG 791
QY      83 AsnGlu-----PhePheAlaPro----- 88
Db      792 GGTGAACATTTACCAATCGAATTCGTGAGCAAAATATTCTTGCCCTCCAGTTCCACAT 851
QY      89 -----AlaSerLeuLeuProAspAla 96
Db      852 CTCATTGAATTGGCTTCAGGCTCGTCGATGACGG 887

RESULT 13
US-09-865-171-26

```



```
; Sequence 26, Application US/09865171
; Patent No. 6451576
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Crock, John E
; APPLICANT: Bohlmann, Jorg
; APPLICANT: Steele, Christopher L
; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS),
; FILE REFERENCE: MSUR117468
; CURRENT APPLICATION NUMBER: US/09/865,171
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/234,393
; PRIOR FILING DATE: 1999-01-20
; PRIOR APPLICATION NUMBER: 60/072,204
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Abies grandis
US-09-865-171-26

Alignment Scores:
Pred. No.: 1.51 Length: 1416
Score: 69.50 Matches: 41
Percent Similarity: 40.91% Conservative: 13
Best Local Similarity: 31.06% Mismatches: 37
Query Match: 13.44% Indels: 41
DB: Gaps: 10

US-10-001-254-6 (1-98) x US-09-865-171-26 (1-1416)
QY 3 ValArgCysLeuAsnValGlyLeuIleArgLysLeuSerAspPheIleAsp---ProGln 21
DB 501 GTGAGACGATGGAGATGTTTCGTGTGTAGAGGCGCTCCACACTTCATGAAATTCATTC 560
QY 22 GluGlyThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 36
DB 561 GAGTTCGTGTAAAGACATCTAATGAATGATGCTGAAGCTGTT---AAAGCGCAAGG 617
QY 37 AspAsp---ArgTyrAsnGlnPheHisIleArgArgPheGluAlaLeu 51
DB 618 CAAGATATGGCGGCTACATAGAAAATGATCGAGG---CGATACCTTGAAAGCTTAT 674
QY 52 LeuGln---ThrGlyLysSerProThrSerGluLeuLeuPheAsp 65
DB 675 CTGCAAGATGGCGGAGATGATAGCCACATGTCCTCCACCTTGATGAGTACTTGAT 734
QY 66 TrpGlyThrThrAsn---CysThrValGlyAspLeuValAspLeuIleGln 82
DB 735 AATGGCACACCAACACTCGGATGTGTATG---AATTGATTCGCGCTTGTGTATG 791
QY 83 AsnGln---PhePheAlaPro----- 88
DB 792 GGTAAACATTACCAATCGACATTCGTGAGCAAAATATTTCTTCCCTCCAGGTTCCACCAT 851
QY 89 -----AlaSerLeuLeuLeuProAspAla 96
DB 852 CTCATTGAATTGGCTTCACAGGCTCGTCGATGACGCG 887

RESULT 14
US-09-234-393-49
; Sequence 49, Application US/09234393A
; Patent No. 6265639
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Crock, John E
; APPLICANT: Bohlmann, Jorg
; APPLICANT: Steele, Christopher L
; APPLICANT: Jetter, Reinhard
; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)
; FILE REFERENCE: MSUR113345
; CURRENT APPLICATION NUMBER: US/09/234,393A
```

```
; TITLE OF INVENTION: AND METHODS OF USE
; FILE REFERENCE: MSUR113345
; CURRENT APPLICATION NUMBER: US/09/234,393A
; EARLIER FILING DATE: 1999-01-20
; EARLIER APPLICATION NUMBER: 60/072,204
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
; LENGTH: 1785
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: gamma humulene synthase
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4)..(1782)
US-09-234-393-49

Alignment Scores:
Pred. No.: 2.11 Length: 1785
Score: 69.50 Matches: 41
Percent Similarity: 40.91% Conservative: 13
Best Local Similarity: 31.06% Mismatches: 37
Query Match: 13.44% Indels: 41
DB: Gaps: 10

US-10-001-254-6 (1-98) x US-09-234-393-49 (1-1785)
QY 3 ValArgCysLeuAsnValGlyLeuIleArgLysLeuSerAspPheIleAsp---ProGln 21
DB 1087 GTGAGACGATGGAGATGTTTCGTGTGTAGAGGCGCTCCACACTTCATGAAATTCATTC 1146
QY 22 GluGlyThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 36
DB 1147 GAGTTCGTGTAAAGACATCTAATGAATGATGCTGAAGCTGTT---AAAGCGCAAGG 1203
QY 37 AspAsp---ArgTyrAsnGlnPheHisIleArgArgPheGluAlaLeu 51
DB 1204 CAAGATATGGCGGCTACATAGAAAATGATCGAGG---CGATACCTTGAAAGCTTAT 1260
QY 52 LeuGln---ThrGlyLysSerProThrSerGluLeuLeuPheAsp 65
DB 1261 CTGCAAGATGGCGGAGATGATAGCCACTGACATGTCCTCCACCTTGATGAGTACTTGAT 1320
QY 66 TrpGlyThrThrAsn---CysThrValGlyAspLeuValAspLeuIleGln 82
DB 1321 AATGGCACACCAACACTCGGATGTGTATG---AATTGATTCGCGCTTGTGTATG 1377
QY 83 AsnGln---PhePheAlaPro----- 88
DB 1378 GGTAAACATTACCAATCGACATTCGTGAGCAAAATATTTCTTCCCTCCAGGTTCCACCAT 1437
QY 89 -----AlaSerLeuLeuLeuProAspAla 96
DB 1438 CTCATTGAATTGGCTTCACAGGCTCGTCGATGACGCG 1473

RESULT 15
US-09-234-393-51
; Sequence 51, Application US/09234393A
; Patent No. 6265639
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Crock, John E
; APPLICANT: Bohlmann, Jorg
; APPLICANT: Steele, Christopher L
; APPLICANT: Jetter, Reinhard
; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)
; FILE REFERENCE: MSUR113345
; CURRENT APPLICATION NUMBER: US/09/234,393A
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 18, 2004, 01:04:18 ; Search time 1658.38 Seconds
(without alignments)
2417.512 Million cell updates/sec

Title: US-10-001-254-6

Perfect score: 517
Sequence: 1 TYVRCINVGIRKLSDFIDP.....LLIIONEFFAPASLLLPDAVP 98

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 segs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n model -DEV=xlh
-Q=/cgr2_1/USPRO.spool/US10001254/runat_16012004_152423_19714/app_query.faeta_1.1109
-DB=GenBdb1 -QPM=faetap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=-1 -MATRIX=Dlosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10001254 @CGN 1.1 5671 @runat_16012004_152423_19714 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenBdb1:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rdd:*
36: em_hcg_mam:*
37: em_hcg_vtc:*
38: em_hcg:*
39: em_hcg_hum:*
40: em_hcg_mus:*
41: em_hcg_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	517	100.0	294	6 AX431296	AX431296 Sequence
2	517	100.0	833	6 AR223870	AR223870 Sequence
3	517	100.0	1383	6 AX431305	AX431305 Sequence
4	517	100.0	1629	6 BC013316	BC013316 Homo sapi
5	517	100.0	1629	6 AX431318	AX431318 Homo sapi
6	517	100.0	2817	6 AF155118	AF155118 Homo sapi
7	517	100.0	2820	9 AK000528	AK000528 Homo sapi
8	513	99.2	1383	6 AX196260	AX196260 Sequence
9	513	99.2	1383	6 AF445802	AF445802 Homo sapi
10	476	92.1	1542	6 AX196262	AX196262 Sequence
11	476	92.1	1542	10 AF445803	AF445803 Sequence
12	476	92.1	2431	10 BC051676	BC051676 Mus muscu
13	413	79.9	501	6 AX321132	AX321132 Sequence
14	287	55.5	1719	6 BC045381	BC045381 Dario rer
15	256	49.5	3303	9 AY186092	AY186092 Homo sapi
16	256	49.5	118572	9 AC093012	AC093012 Homo sapi
17	256	49.5	185868	2 AC021719	AC021719 Homo sapi
18	256	49.5	309582	2 AC025567	AC025567 Homo sapi
19	247	47.8	2213	6 AX431316	AX431316 Sequence
20	247	47.8	2213	6 BD155790	BD155790 Primer fo
21	247	47.8	2213	9 AK027301	AK027301 Homo sapi
22	233	45.1	264657	2 AC129390	AC129390 Rattus no
23	123	23.8	756	11 BV075713	BV075713 S212P6044
24	121.5	23.5	2069	3 AK116443	AK116443 Clona int
25	98.5	19.1	1830	10 MM0440757	MM0440757 Mus muscu
26	96.5	18.7	1791	6 AX056430	AX056430 Sequence
27	96.5	18.7	2277	6 AX057324	AX057324 Sequence
28	96.5	18.7	2288	9 AF113136	AF113136 Homo sapi
29	95	18.4	1888	10 AF461763	AF461763 Mus muscu
30	88	17.0	161894	2 AC137339	AC137339 Rattus no
31	88	17.0	249033	2 AC107528	AC107528 Rattus no
32	86.5	16.7	213254	2 AC149424	AC149424 Mus muscu
33	86.5	16.7	244686	2 AC123081	AC123081 Rattus no
34	86.5	16.7	252400	2 AC094859	AC094859 Rattus no
35	86	16.6	2603	5 BC043819	BC043819 Xenopus 1
36	84	16.2	125777	9 AC078927	AC078927 Homo sapi
37	84	16.2	207951	2 AC012264	AC012264 Homo sapi
38	84	16.2	279125	2 AC113694	AC113694 Rattus no
39	84	16.2	288358	2 AC106121	AC106121 Rattus no
40	82	15.9	1947	10 BC005591	BC005591 Mus muscu
41	82	15.9	1960	10 BC051075	BC051075 Mus muscu
42	82	15.9	2606	10 MM089023	MM089023 Mus muscu
43	82	15.9	159431	10 AC128702	AC128702 Mus muscu
44	82	15.9	177210	2 AC134406	AC134406 Mus muscu
45	81	15.7	143452	2 AC136209	AC136209 Gallus ga

RESULT 1

ALIGNMENTS

TITLE	Knuth,A.
JOURNAL	Renal cancer associated antigens and uses therefor
PATENT	Patent: US 640663-A 10 27-AUG-2002;
FEATURES	Location/Qualifiers
SOURCE	1. 833 /organism="unknown"
BASE COUNT	273 a 155 c 179 g 226 t
ORIGIN	
Alignment Scores:	
Prod. No.:	3.92e-61 Length: 833
Score:	517.00 Matches: 98
Percent Similarity:	100.00% Conservative: 0
Best local similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	Gaps: 0
US-10-001-254-6 (1-98) x AR223870 (1-833)	
QY	1 ThrTyrValArgCysLeuAsnValGlyLeuIleAlrghysLeuSerAaspheileAaspPro 20
Db	74 ACATAGTGCGCCGCCCAATGGTTGACATAATTAGGAAGCTGTGCAGATTATTATGACTCT 133
QY	21 GlnGluGlyTrpPylsLysLeuAlavalAlalelYslYsPserOsergiYaasParqTYr 40
Db	134 CAGAAGCATGGAAAGAAAGTAGTACGTGTACATTAATAAACCATCTGGTGTAGATAGATAC 193
QY	41 AsnGlnPheHisIleArgArgPheGlnAlaleuLeuGlnThrGlyLysePProThSer 60
Db	194 AATCAGTTTCACATTAAGAGATTGGAAGCATTTACTTCMAACTGAAAAAGTCCACACTTCT 253
QY	61 GluIleuLeuPheAspTrpGlyThrThrAsnCysThrValGlyAspLeuValAspLeuLeu 80
Db	254 GAATTACTGTTTGGCTGGGCGCACCAAAATGGACAAGTGGTGAATCTTGGATCTTTTG 313
QY	81 IleGlnAsnGluPhePheAlaProAlaSerLeuLeuLeuProAspaAlavalPro 98
Db	314 ATCCAAAATGAATTTTTTGTCTCTCGCAGACTTTTGTCTCCAGATGCTGTTCCC 367
RESULT 3	
AX431306	1383 bp DNA linear PAT 28-JUN-2002
LOCUS	Sequence 15 from Patent WO0240660.
ACCESSION	AX431306
VERSION	AX431306.1 GI:21656175
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 Pawlowski,K., Fiorentino,L., Godzik,A., Lee,S.H., Reed,J.C., Roth,W. and Stenmer-Liewen,F.
TITLE	Novel death domain proteins
JOURNAL	Patent: WO 0240680-A 15 23-MAY-2002;
FEATURES	BURNHAM INST (US)
SOURCE	Location/Qualifiers
	1. 1383
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	1. 1383
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAD37279.1"
	/db_xref="GI:21656176"
CDS	/translation="VANKPTIPSTYYRCLANGRLKSDPIIDPGMKKLAVAIKKPSGDDYNQHIRRFALLQTGSPTSSELPDWTNCTCYGDVLILQNEFRPASLLFDPVKANTLPSEKAITVQQKMPFCDKDKTLMTPVNLQSSVMPPOSSSPENSLEVDPRHFHSFEYLKNVTNNNDERPI SVGNKMGGGGGVYKYGVNTTVAAVKLIAMVDLTTELTKOOPDEIKVMACCOHEMLVELLGSSDDGDCLVVVWPNGLIDRLSLGIDTPPTLSMWRCKKIAGANGINTEHLHHIRDISANILLDEAFTRAKISDGLARASEFPATVTSRTSVGITAMAPALKEGITPPSDITSFGVILEITITGLPAVDENRE

CDS

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
50..1432
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD37285.1"
/db_xref="GI:21656188"

/translation="MNKPTSTYVRCINVLGRLKSLDFIDPEGKMLAVAKKPSG
DRLVNFHRRPEALLOTGKSPSELPWMCTNCTVGLVDLLIQNEFPASILLP
DAVPTANTLPSKEALTVQOKMPFCDKRTTLTPVQNI.EOSTMPDSSSPKSLFV
SDTRHSPFELKVNTPNDFRPIVSGNKGEGFVGVYKGVNNTTAAVKILAM
VDITTEELKQOFDOI.KYMAKCOHENLVELLFGSSGDGLCLVYVMPGSLDLRLSC
LDGTPPLSMHMRCKIAQGAANGINFLHNHHRIDIKSANILLDEAFATKISDFGLAR
ASEKFAQVWTSRIVGTITAYMAPEALRGETTPKSDIYSGVVLLEITITGLPAVDHRE
POLLIDIKKEIEDEKTIEDYIDKKNDADSTSVAMYSAGSQCCHKKKNSPDIKV
HOLLQEMTAS"

BASE COUNT 912 a 547 c 586 g 772 t
ORIGIN

Alignment Scores:

Pred. No.: 1,71e-60 Length: 2817
Score: 517.00 Matches: 98
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-10-001-254-6 (1-98) x AX431318 (1-2817)

QY 1 ThrryValArgCysLeuAsnValGlyLeuIleArgIysLeuSerAspPheIleAspPro 20
Db 74 ACATATGCGGCGCTCCATGTTGACATTAATAGAGAGCTGCAGATTTTATGATTCCT 133
QY 21 GInGluGlyTrpIysLeuAlaValAlaIleIysLysProSerGlyAspAspArgTyr 40
Db 134 CAGAAGAGTGAAGAAGATTAGCTAGCTATTAAAAACATCTGGTGTATGATGATAC 193
QY 41 AAsnGlnPheHisIleArgArgPheGluAlaLeuLeuGlnThrGlyLysSerProThrSer 60
Db 194 AATCAGTTTACATTAAGAGATTGAAACATTAATCTGAACTGAAAGATCCCACTTCT 253
QY 61 GluLeuLeuPheAspTrpGlyThrThrAsnCysThrValGlyAspLeuValAspLeu 80
Db 254 GAATTAAGTGTGACAGCGGACCAACAATTTGACAGTTCGTGATCTTGTGATCTTTTG 313
QY 81 ILeGlnAsnGluPheAlaProIleAsrLeuLeuProAspAlaValPro 98
Db 314 ATCCAAATGAATTTTTCCTCTCGAGAGCTTTTGTCTCCAGATGCTGTTCCC 367

RESULT 6
AP155118 2817 bp mRNA linear PRI 05-JAN-2000
LOCUS Homo sapiens putative protein kinase NY-REN-64 antigen mRNA,
DEFINITION complete cds.
ACCESSION AP155118
VERSION AP155118.1 GI:5360130
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2817)
Scanlan,M.J., Gordan,J.D., Williamson,B., Stockert,E., Bander,N.H.,
Jongeneel,V., Gure,A.O., Jager,D., Jager,E., Knuth,A., Chen,Y.-T.
and Old,L.J.
Antigens recognized by autologous antibody in patients with
renal-cell carcinoma
Int. J. Cancer 83 (4), 456-464 (1999)
JOURNAL 99438124
MEDLINE 10508479
PUBMED 2 (bases 1 to 2817)
REFERENCE Scanlan,M.J., Gordan,J.D., Williamson,B., Stockert,E., Bander,N.H.,

Jongeneel,V., Gure,A.O., Jager,D., Jager,E., Knuth,A., Chen,Y.-T.
and Old,L.J.
Direct Submission
JOURNAL Submitted (28-MAY-1999) Ludwig Institute, Sloan-Kettering
Institute, 1275 York Ave, New York, NY 10021, USA
FEATURES
source
1..2817
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="1973/10-4"
/cell_type="renal cell carcinoma"
50..1432
/codon_start=1
/product="putative protein kinase NY-REN-64 antigen"
/protein_id="AAD4284.1"
/db_xref="GI:5360131"

CDS

/translation="MNKPTSTYVRCINVLGRLKSLDFIDPEGKMLAVAKKPSG
DRLVNFHRRPEALLOTGKSPSELPWMCTNCTVGLVDLLIQNEFPASILLP
DAVPTANTLPSKEALTVQOKMPFCDKRTTLTPVQNI.EOSTMPDSSSPKSLFV
SDTRHSPFELKVNTPNDFRPIVSGNKGEGFVGVYKGVNNTTAAVKILAM
VDITTEELKQOFDOI.KYMAKCOHENLVELLFGSSGDGLCLVYVMPGSLDLRLSC
LDGTPPLSMHMRCKIAQGAANGINFLHNHHRIDIKSANILLDEAFATKISDFGLAR
ASEKFAQVWTSRIVGTITAYMAPEALRGETTPKSDIYSGVVLLEITITGLPAVDHRE
POLLIDIKKEIEDEKTIEDYIDKKNDADSTSVAMYSAGSQCCHKKKNSPDIKV
HOLLQEMTAS"

BASE COUNT 912 a 547 c 586 g 772 t
ORIGIN

Alignment Scores:

Pred. No.: 1,71e-60 Length: 2817
Score: 517.00 Matches: 98
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-10-001-254-6 (1-98) x AP155118 (1-2817)

QY 1 ThrryValArgCysLeuAsnValGlyLeuIleArgIysLeuSerAspPheIleAspPro 20
Db 74 ACATATGCGGCGCTCCATGTTGACATTAATAGAGAGCTGCAGATTTTATGATTCCT 133
QY 21 GInGluGlyTrpIysLeuAlaValAlaIleIysLysProSerGlyAspAspArgTyr 40
Db 134 CAGAAGAGTGAAGAAGATTAGCTAGCTATTAAAAACATCTGGTGTATGATGATAC 193
QY 41 AAsnGlnPheHisIleArgArgPheGluAlaLeuLeuGlnThrGlyLysSerProThrSer 60
Db 194 AATCAGTTTACATTAAGAGATTGAAACATTAATCTGAACTGAAAGATCCCACTTCT 253
QY 61 GluLeuLeuPheAspTrpGlyThrThrAsnCysThrValGlyAspLeuValAspLeu 80
Db 254 GAATTAAGTGTGACAGCGGACCAACAATTTGACAGTTCGTGATCTTGTGATCTTTTG 313
QY 81 ILeGlnAsnGluPheAlaProIleAsrLeuLeuProAspAlaValPro 98
Db 314 ATCCAAATGAATTTTTCCTCTCGAGAGCTTTTGTCTCCAGATGCTGTTCCC 367

DNA encoding sea ID #6

RESULT 7
AK000528 2820 bp mRNA linear PRI 22-FEB-2000
LOCUS Homo sapiens CDNA FLJ20521 f1f, clone KAT10395.
DEFINITION
ACCESSION AK000528
VERSION AK000528.1 GI:7020663
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2820)
Scanlan,M.J., Gordan,J.D., Williamson,B., Stockert,E., Bander,N.H.,
Jongeneel,V., Gure,A.O., Jager,D., Jager,E., Knuth,A., Chen,Y.-T.
and Old,L.J.
Antigens recognized by autologous antibody in patients with
renal-cell carcinoma
Int. J. Cancer 83 (4), 456-464 (1999)
JOURNAL 99438124
MEDLINE 10508479
PUBMED 2 (bases 1 to 2817)
REFERENCE Scanlan,M.J., Gordan,J.D., Williamson,B., Stockert,E., Bander,N.H.,

Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
Ota,T., Suzuki,Y., Ohashi,M., Nishi,T., Shibahara,T., Tanaka,T.,

TITLE Nakamura, Y., Isogai, T. and Sugano, S.
JOURNAL NEDO human cDNA sequencing project
REFERENCE 2 (bases 1 to 2820)
AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obaashi, M., Nishi, T., Isogai, T., Shibanari, T., Tanaka, T. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-3286, Fax: 81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

Source 1..2820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KAT10395"
/cell_line="KATO III"
/cell_type="signet-ring cell carcinoma"
/clone_lib="KAT"
/note="Cloning vector pME18SFL3"
30..1412
/note="unnamed protein product"
/codon_start=1
/protein_id="BA91232.1"
/db_xref="GI:7020684"
/translation="MNKPTPTSTYVRCNVGLIRKLSDFIDPOEGMKLVAIAIKRPSG
DVRVNFHIRFEPALLOTGKSPSELFPWGTNTCVGLDVLILQNEFPAPSLILP
DAVPRKANTLPSKEALTIVQOKMPFCOKRITLTPVQNLQSQSYMPDSSPEKSLIEV
SDTRFHSFSEYELKNTNFDERPI SVGNKMGEGGVYKGVNNTVAVKKLAM
VDITTEELKQFPOEIKVMAKCOHENLVELLGFSSDGLCLVYVMPGSLDLRLSC
LDGPTPLSMHMKCKIAQGAANGINFLENHHRIDIKSANIILDEAFYAKISDFGLAR
ASEKFAQTWTSRIVGTATYMAPEALRGETTPSDIYSFGVVLLEITIGLPAYDERE
POLLDIKKEIDEKTEIDYIDKKNADDSVEAMYSVAQCLHEKKNRDIDIKV
QQLDQMTAS"

CDS

BASE COUNT 940 a 534 c 579 g 767 t
ORIGIN

Alignment Scores:
Pred. No.: 1,71e-60 Length: 2820
Score: 517.00 Matches: 98
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-001-254-6 (1-98) x AK000528 (1-2820)

QY 1 ThrTyrValArgCysLeuAsnValAGlyLeuIleArgLysLeuSerAspPheIleaspPro 20
DB 54 ACATATGTGGCGCTGCCATGTTGACTAATTHGAGAGCTGTCAGATTATTATGATCCT 113
QY 21 GlnGlnGlyTyrPylsYsLeuAlaValAlaIleLysLysProSerGlyAspAspArgTyr 40
DB 114 CAGAAGAGATGAGAAGATTAGCTGATGATTAAAAACCATCTGCTGATGATAGTAC 173
QY 41 AenGlnPheHisIleArgArgPheGlnAlaLeuLeuGlnThrGlyLysSerProThrSer 60
DB 174 AATCAGTTTCACATTAAGAGATTGGAAGCATTTCTCAAACTGGAAAAAGTCCACATCTT 233
QY 61 GluLeuLeuPheAspTyrPylsYsLeuAlaValAlaIleLysLysProSerGlyAspAspArgTyr 80
DB 234 GAATTACTGTTTGACTGGGACACCAAAATTGACAGCTTGATCTTGTGATCTTTTG 293
QY 81 IIEGlnAsnGlnPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
|||||

Db 294 ATCCAAATGAAATTTTGTCTCTGCGAGCTTTTGTCTCCAGATGCTGTCC 347

RESULT 8
LOCUS AX196260 1383 bp DNA linear PAT 28-AUG-2001
DEFINITION Sequence 2 from Patent WO0151641.
ACCESSION AX196260
VERSION AX196260.1 GI:15386462
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Wesche, H. and Li, S.
Irak-4: compositions and methods of use
Patent: WO 0151641-A 2 19-JUL-2001;
Tularik Inc. (US)

FEATURES

Source 1..1383
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="human IL-1 receptor-associated kinase 4 (IRAK-4)
cDNA"

CDS

1..1383
/note="human IRAK-4"
/codon_start=1
/protein_id="CAC60090.1"
/db_xref="GI:15386463"
/translation="MNKPTPTSTYVRCNVGLIRKLSDFIDPOEGMKLVAIAIKRPSG
DVRVNFHIRFEPALLOTGKSPSELFPWGTNTCVGLDVLILQNEFPAPSLILP
DAVPRKANTLPSKEALTIVQOKMPFCOKRITLTPVQNLQSQSYMPDSSPEKSLIEV
SDTRFHSFSEYELKNTNFDERPI SVGNKMGEGGVYKGVNNTVAVKKLAM
VDITTEELKQFPOEIKVMAKCOHENLVELLGFSSDGLCLVYVMPGSLDLRLSC
LDGPTPLSMHMKCKIAQGAANGINFLENHHRIDIKSANIILDEAFYAKISDFGLAR
ASEKFAQTWTSRIVGTATYMAPEALRGETTPSDIYSFGVVLLEITIGLPAYDERE
POLLDIKKEIDEKTEIDYIDKKNADDSVEAMYSVAQCLHEKKNRDIDIKV
QQLDQMTAS"

BASE COUNT 463 a 244 c 283 g 393 t
ORIGIN

Alignment Scores:
Pred. No.: 2,56e-60 Length: 1383
Score: 513.00 Matches: 97
Percent Similarity: 98.98% Conservative: 0
Best Local Similarity: 98.98% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 6 Gaps: 0

US-10-001-254-6 (1-98) x AX196260 (1-1383)

QY 1 ThrTyrValArgCysLeuAsnValAGlyLeuIleArgLysLeuSerAspPheIleaspPro 20
DB 25 ACATATGTGGCGCTGCCATGTTGACTAATTHGAGAGCTGTCAGATTATTATGATCCT 84
QY 21 GlnGlnGlyTyrPylsYsLeuAlaValAlaIleLysLysProSerGlyAspAspArgTyr 40
DB 85 CAGAAGAGATGAGAAGATTAGCTGATGATTAAAAACCATCTGCTGATGATAGTAC 144
QY 41 AenGlnPheHisIleArgArgPheGlnAlaLeuLeuGlnThrGlyLysSerProThrSer 60
DB 145 AATCAGTTTCACATTAAGAGATTGGAAGCATTTCTCAAACTGGAAAAAGTCCACATCTT 204
QY 61 GluLeuLeuPheAspTyrPylsYsLeuAlaValAlaIleLysLysProSerGlyAspAspArgTyr 80
DB 205 GAATTACTGTTTGACTGGGACACCAAAATTGACAGCTGATCTTGTGATCTTTTG 264
QY 81 IIEGlnAsnGlnPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
DB 265 ATCCAAATGAAATTTTGTCTCTGCGAGCTTTTGTCTCCAGATGCTGTCC 318
|||||

RESULT 9

AF445802
LOCUS AF445802 1383 bp mRNA linear PRI 20-APR-2002
DEFINITION Homo sapiens interleukin-1 receptor associated kinase 4 (IRAK4)
ACCESSION AF445802
VERSION AF445802.1 GI:20219009
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Li, S., Strelow, A., Fontana, E.J. and Wesche, H.
TITLE IRAK-4: a novel member of the IRAK family with the properties of an IRAK-kinase
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (8), 5567-5572 (2002)
MEDLINE 21957277
PUBMED 11960013
REFERENCE
AUTHORS Suzuki, N., Suzuki, S., Duncan, G.S., Millar, D.G., Wada, T., Mitsuo, C., Takada, H., Wakeham, A., Ito, A., Li, S., Penninger, J.M., Wesche, H., Ohashi, P.S., Mak, T.W. and Yeh, W.C.
TITLE Severe impairment of interleukin-1 and Toll-like receptor signaling in mice lacking IRAK-4
JOURNAL Nature 416 (6882), 750-756 (2002)
MEDLINE 11923871
PUBMED
REFERENCE
AUTHORS Li, S., Strelow, A., Fontana, E.J. and Wesche, H.
TITLE Direct Submision
JOURNAL Submitted (08-NOV-2001) Biology 1, Tularik Inc., 2 Corporate Drive, South San Francisco, CA 94080, USA
FEATURES
source Location/Qualifiers
1..1383
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1383
/gene="IRAK4"
1..1383
/gene="IRAK4"
/note="protein kinase"
/codon_start=1
/product="interleukin-1 receptor associated kinase 4"
/protein_id="AA015772.1"
/db_xref="GI:20219010"
/translation="MNKPTTPSTYTYRCLNVLGRKLSDFIDPOEGMKLAIAIKKPSG
DVRVQPHIRREPEALLOTGKSPTELLFDWGTTCCTAGDLVLLIQNEFFAPSLILP
DAVYKNTLPSKEALTVOQKMPFCDKRTMTPTVQNLQOSYMPDSSSPNKSLIEV
SDTRFHSFSEYELKNTNNNDERPIISVGNKMGEGFVYKGVYNNITVAIVKLAAM
VDITTEELKQGFPOEIKVNAKCOHENLVELGFSDDDLCLVYVYMPGSLDLRLSC
LGCTPPLSMHMRCKIAQGAANGINPLHNNHHRDIKSNILLDKDFAKISDFGLAR
ASERKQVMTSRIVGTITAYMAPEALRGELTPKSDIYSFGVVLLELITGLPAVDENRE
POLLDIKEIEDEBEKTIEDYTDKKNVDADSDVEAMYSVAASQCLHEKKNRPRDIKVV
QQLQEMTAS"

BASE COUNT 463 a 244 c 283 g 393 t
ORIGIN
Alignment Scores:
Pred. No.: 2,56e-60 Length: 1383
Score: 513.00 Matches: 97
Percent Similarity: 98.98% Conservative: 0
Best Local Similarity: 98.98% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: Gaps: 0
US-10-001-254-6 (1-98) x AF445802 (1-1383)
QY 1 ThTtYValArGcYsleuAnVaIglYleuIlleArGlySleuSerAspPheIlleAspPro 20
Db 25 ACATATGCTGGCGCTCAATGTTGACTTAATAGGAAGCTGTCAAGATTTTATTTGATCTT 84
QY 21 GInGluGlyTTrpLySylsleuAlaValAlaIlelySylsPProSerGlyAspAspArgTyR 40

|||||
Db 85 CAAGAGGATGGAAGAACTAGCTGATCTTTAAAAACCATCTGGAGATAGATAC 144
QY 41 AenGlnPheHsIleArGArGPhGluAlaIleuEngInThnGlylVsSerProThrSer 60
Db 145 AATCAGTTTCACATDAGAGATTGAGCATTTACTTCAACTGGAAAAAGTCCACCTTCT 204
QY 61 GluLeuLeuPheAspTrpGlyThrThrasnCystrValGlyAspLeuValAspLeuLeu 80
Db 205 GAATTACTGTTTGAATGGGCGACACCAAAATTGCACAGCTGGATCTTGATCTTTTG 264
QY 81 lIleGlnAsnGluPhePheAlaProAlaSerIleuLeuPProAspAlaValPro 98
Db 265 ATCCAAATGAATTTTTCCTGCTGCGAGTCTTTTGTCCAGATGCTGTTCCC 318
RESULT 10
LOCUS AX196262 1542 bp DNA linear PAT 28-AUG-2001
DEFINITION Sequence 4 from Patent WO0151641.
ACCESSION AX196262
VERSION AX196262.1 GI:15386464
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Wesche, H. and Li, S.
TITLE IRAK-4: compositions and methods of use
JOURNAL Patent: WO 0151641-A 4 19-JUL-2001;
Tularik Inc. (US)
FEATURES
source Location/Qualifiers
1..1542
/organism="Mus sp."
/mol_type="genomic DNA"
/db_xref="taxon:10095"
/note="murine IL-1 receptor-associated kinase 4 (IRAK-4)
cDNA"
153..1542
/note="murine IRAK-4"
/codon_start=1
/protein_id="CAC60091.1"
/db_xref="GI:15386465"
/translation="MNKPLTPSTYTYRINLVNGLRKLSDPIDPOEGMKLAIAIKKPSG
DVRVQPHIRREPEALLOTGKSPTELLFDWGTTCCTAGDLVLLIQNEFFAPSLILP
DAVYKNTLPSKEALTVOQKMPFCDKRTMTPTVQNLQOSYMPDSSSPNKSLIEV
SDTRFHSFSEYELKNTNNNDERPIISVGNKMGEGFVYKGVYNNITVAIVKLAAM
VDITTEELKQGFPOEIKVNAKCOHENLVELGFSDDDLCLVYVYMPGSLDLRLSC
LGCTPPLSMHMRCKIAQGAANGINPLHNNHHRDIKSNILLDKDFAKISDFGLAR
ASERKQVMTSRIVGTITAYMAPEALRGELTPKSDIYSFGVVLLELITGLPAVDENRE
POLLDIKEIEDEBEKTIEDYTDKMSDADSDASVEAMYSVAASQCLHEKKNRPRDIKVV
QQLQEMSA"

BASE COUNT 421 a 392 c 423 g 306 t
ORIGIN
Alignment Scores:
Pred. No.: 3.5e-55 Length: 1542
Score: 476.00 Matches: 89
Percent Similarity: 95.92% Conservative: 5
Best Local Similarity: 90.82% Mismatches: 4
Query Match: 92.07% Indels: 0
DB: Gaps: 0
US-10-001-254-6 (1-98) x AX196262 (1-1542)
QY 1 ThTtYValArGcYsleuAnVaIglYleuIlleArGlySleuSerAspPheIlleAspPro 20
Db 187 ACATATGCTGGCGCTCAATGTTGACTTAATAGGAAGCTGTCAAGATTTTATTTGATCTT 246
QY 21 GInGluGlyTTrpLySylsleuAlaValAlaIlelySylsPProSerGlyAspAspArgTyR 40
Db 247 CAAGAGGATGGAAGAAATTTAGCATGATCAAAAAGCGCTCCGCGACAGACATAC 306

Qy 41 AsnGlnPheHisIleArgArgPheGluAlaIleuGlnThrGlyIleSerProThSer 60
Db 307 AATCAGTTCATATTAAGAGATTCGAGACCTTACTTACAGACCGGAGAGACCCACCTGT 366
Qy 61 GluIleuPheAspTrpGlyThrThrAsnCyThrValGlyAspIleValAspIleu 80
Db 367 GAACGTGCTGTGACTGGGGCAACGACGACGACGAGTGGCGACTTGTGATCTACTG 426
Qy 81 IleGlnAsnGluPhePheAlaProAlaSerIleuIleuProAspAlaValPro 98
Db 427 GTCAGATTGAGCTGTTGGCCCCGCCACCTCTCTGCTGGGATGCCGTTCCC 480

RESULT 11
AF445803 1542 bp mRNA linear ROD 20-APR-2002
LOCUS Mus musculus interleukin-1 receptor associated kinase 4 (Itak4)
DEFINITION mRNA, complete cds.
ACCESSION AF445803.1 GI:20219011
VERSION
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM

REFERENCE 1 (bases 1 to 1542)
AUTHORS Li, S., Strelow, A., Fontana, E. J. and Wesche, H.
TITLE IRAK-4: a novel member of the IRAK family with the properties of an IRAK-kinase
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (8), 5567-5572 (2002)
MEDLINE 21957277
PUBMED 11960013
REFERENCE 2 (bases 1 to 1542)
AUTHORS Li, S., Strelow, A., Fontana, E. J. and Wesche, H.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-2001) Biology I, Tularek Inc., 2 Corporate Drive, South San Francisco, CA 94080, USA

FEATURES
Source
1. 1542
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
1. 1542
/gene="Itak4"
163. 1542
/gene="Itak4"
/note="protein kinase"
/codon_start=1
/product="interleukin-1 receptor associated kinase 4"
/protein_id="AA15773.1"
/db_xref="GI:20219012"
/translation="MNKPLTPSTYIRNLANGIRKISDPIDPOGKMLAVALIKKSG
DDRNRPHIRFELQTKSPTELLFDGTTNCTVGDVLVLEIAPATLLP
DAVQIVKSLPRPAAVQAQTHGCEKQKTSWMPKLEHSCBPSSPDRSVES
SDRFHSFSEHLEKSIITNDEQPASAGRMGGGVVKGCVNNTIYAVKLGAM
VEISTELKCOFPOEIKVMATCOHENLVELLGGSSDNLCLVYAWPNSSILDRSC
LDGTPPLSMHTRCKVAGCTANGIRFLHSHHHRDLSANILLDPKTAISDGLAR
ASARLAQTWTSRIKVTGTAIAPALPAGELTTPSDITSFVAILLITGLAAVDENR
POLLLDKEIIDEKTEIDYDEKMSDADPASVEANSAQCLHEKKRRPDIAYK
QQLLOEMSA"

BASE COUNT 421 a 392 c 423 g 306 t
ORIGIN
Alignment Scores:
Pred. No.: 3.5e-55 Length: 1542
Score: 476.00 Matches: 89
Percent Similarity: 95.92% Conservative: 5
Best Local Similarity: 90.82% Mismatches: 4
Query Match: 92.07% Indels: 0
DB: 10 Gaps: 0

US-10-001-254-6 (1-98) x AF445803 (1-1542)

Qy 1 ThrTyValArgCysIleuAsnValGlyIleuIleArgIleuSerAspPheIleAspPro 20
Db 187 ACATACCTACGCAACCTTATGCGGATCTTATAGAGAGCTGTGATTTATGATCT 246
Qy 21 GlnGluGlyTrpIleValIleValAlaIleIleValIleProSerGlyAspAspArgTyr 40
Db 247 CAAGAAGGGTGAAGAAATTAAGAGTAGATCATCAAAAAGCCGCGGACGACAGATAC 306
Qy 41 AsnGlnPheHisIleArgArgPheGluAlaIleuGlnThrGlyIleSerProThSer 60
Db 307 AATCAGTTCATATTAAGAGATTCGAGACCTTACTTACAGACCGGAGAGACCCACCTGT 366
Qy 61 GluIleuPheAspTrpGlyThrThrAsnCyThrValGlyAspIleValAspIleu 80
Db 367 GAACGTGCTGTGACTGGGGCAACGACGACGACGAGTGGCGACTTGTGATCTACTG 426
Qy 81 IleGlnAsnGluPhePheAlaProAlaSerIleuIleuProAspAlaValPro 98
Db 427 GTCAGATTGAGCTGTTGGCCCCGCCACCTCTCTGCTGGGATGCCGTTCCC 480

RESULT 12
BC051676 2431 bp mRNA linear ROD 10-JUN-2003
LOCUS Mus musculus interleukin-1 receptor-associated kinase 4, transcript variant 1, mRNA (CDNA clone MGC:60994 IMAGE:30017484), complete cds.
ACCESSION BC051676 GI:30354527
VERSION
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM

REFERENCE 1 (bases 1 to 2431)
AUTHORS Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, F. S., Wagner, C. M., Schuler, G. D., Altshuler, S. F., Zeeberg, B., Buettow, K. H., Scheffer, C. F., Bhat, N. K., Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L., Stapleton, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L., Scheetz, T. E., Brownstein, M. J., Useth, T. B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S. S., Loquellano, N. A., Peters, G. J., Abramson, R. D., Mullany, S. J., Bosak, S. A., McEwan, P. J., McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S., Morley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W., Villalón, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shvachenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butlerfield, Y. S., Kravinsky, M. I., Skalska, U., Smal, J., D. E., Scherch, A., Schein, J. E., Jones, S. J. and Marra, M. A.
TITLE human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12479932
REFERENCE 2 (bases 1 to 2431)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (30-APR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Minoru Ko
CDNA Library Preparation: Yulan Piao and Minoru Ko
CDNA Library Arrayed by: The I.W.A.G.E. Consortium (LINL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huljk, S.W., Louieged, H.,
 Kowals, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>
 Series: IRAC Plate: 111 Row: 1 Column: 20
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 23943897.

FEATURES

source

1. 2431
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="MGC:60994 IMAGE:30017484"
 /issue_type="Embryo, whole, mouse, 7.5-dpc"
 /clone_lib="NIA Mouse 7.5-dpc Whole Embryo cDNA Library
 (long)"
 /lab_host="DH10B"
 /note="Vector: pSPORT1"
 1. 2431
 /gene="Irk4"
 /note="synonyms: NY-REN-64, 933020D03R1K"
 /db_xref="locusid:26632"
 /db_xref="MGI:2182474"
 168. 1547
 /codon_start=1
 /product="interleukin-1 receptor-associated kinase 4,
 isoform a"
 /protein_id="AAH51676.1"
 /db_xref="GI:30354528"
 /db_xref="locusid:26632"
 /translacion="NMPLPTSTYIRNLNVGLIRKSDPIDQEWKLAIVAKPESG
 DRYNQFHIRPEALIQTKSPCTCLPMTGTCVGLVDLVOIEFAPATLLP
 DAVPTVKSLPREAATVAQTGHPCEKRTSVMPKLEHSCPEPSSPNRSES
 SDTRFSPFHELKSTINNPDEQPASGNGRMEGGGVYVGCVVNTTVAVKIGAM
 VETSTELKQPDQETKQWATQCHENLVLLGSSPSDNLCLVYATMPGSLDLSC
 LQSTPPLSWHTRKVAQGTANGIRFLHEHNIHRDTSANILDRDFTKISDFGLAR
 ASARLAQVWTRISIVGTAYMAEALRGETTPKSDIYSFGVILELITGLAANDERE
 POLLIDKEIEDEEKTIEDYDEKMSDADPASVEAMSAASQCLEKQNRPRDAKY
 OQLLOEMSA"

CDS

BASE COUNT 656 a 580 c 618 g 577 t
 ORIGIN

Alignment Scores:

Pred. No.: 6.08e-55 Length: 2431
 Score: 476.00 Matches: 89
 Percent Similarity: 95.92% Conservative: 5
 Best Local Similarity: 90.82% Mismatches: 4
 Query Match: 92.07% Indels: 0
 Gaps: 0

US-10-001-254-6 (1-98) x BC051676 (1-2431)

QY 1 ThrTyrValArgCysLeuAenValIGlyLeuLeuLeuLeuSerAspPheIleAapPro 20
 DB 192 ACATACATACGCAACCTTAATGCGGATCTTACGAGAGCTGCGATTATTTATGATCCT 251
 QY 21 GInGluG1YTrpYsLysLeuAlaValAlaIleLysLysProSerG1YAspAspArgTyR 40
 DB 252 CAGAAAGGGTGGAGAAATTAAGAGTACATCAAAAAGCCGTCGGCGAGCAGATAC 311
 QY 41 AAsGlnPheHisIleArgArgPheGluAlaLeuLeuGlnThrG1YysSerProThrSer 60
 DB 312 AATCAGTTCAATTAAGAGATTCGAAGCCTTACTCAGACCGGAGAGAGCCCACTCT 371
 QY 61 GluLeuLeuPheAspTrpG1YThrThrAsnCysThrValG1YspleuValaApleuLeu 80
 DB 372 GAACGTCGTGTTGATCGGGGACACAGAACTGACAGATTGGCGACCTTGTGATCTACTG 431

QY 81 IlegInaAngluPheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
 DB 432 GTCCAGATTGAGCTGTGCTCCCGCCGCACTCTCCTGCTGCGGATGCCGTTCCC 485

RESULT 13

AX321132 501 bp DNA linear PAT 15-DEC-2001
 LOCUS
 DEFINITION Sequence 149 from Patent WO0177168.
 ACCESSION AX321132
 VERSION AX321132.1 GI:17904576
 KEYWORDS

SOURCE

ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Lodes, M.J., Wang, T., Mohamath, R. and Indritas, C.Y.
 Compositions and methods for the therapy and diagnosis of lung
 cancer
 Patent: WO 0177168-A 149 18-OCT-2001;
 CORIXA CORPORATION (US)

JOURNAL

FEATURES

source

1. 501
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 166 a 102 c 97 g 134 t 2 others
 ORIGIN

Alignment Scores:

Pred. No.: 4e-47 Length: 501
 Score: 413.00 Matches: 93
 Percent Similarity: 94.00% Conservative: 1
 Best Local Similarity: 93.00% Mismatches: 4
 Query Match: 79.88% Indels: 4
 Gaps: 0

US-10-001-254-6 (1-98) x AX321132 (1-501)

QY 1 ThrTyrValArgCysLeuAenValIGlyLeuLeuLeuLeuSerAspPheIleAapPro 20
 DB 45 ACATATGCGCTGCTCAATGTTGACCTAATTAAGAGCTGTCAGATTATTTATGATCCT 104
 QY 21 GInGluG1YTrpYsLysLeuAlaValAlaIleLysLysProSerG1YAspAspArgTyR 40
 DB 105 CAGAAAGATGAGAAAGTACGTAGCTATTAATAAACATCTGGGATCATGATAC 164
 QY 41 AAsGln-PheHisIleArgArgPheGluAlaLeuLeuGlnThrG1YysSerProThrSe 60
 DB 165 AATCAAGTTTCATACATAGAGAGATTGAAGCATT-CTTCAAACTGAGAAAAGTCCCACTTC 223
 QY 60 rGluLeuLeuPheAspTrp-G1YThrThrAsnCysThrValG1YspleuValaApleuLeu 80
 DB 224 TTGAATACCTGTTGACTGGGGGACACCAAAATGACAGTTGGTATCTTGTGGATCTTT 283
 QY 80 euILeGlnaAngluPheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
 DB 284 TGAATCAAAATTAATT-TTGTCTCTCGAGAGCTTTGTCTCCCAATGCTGTTC 338

RESULT 14

BC045381 1719 bp mRNA linear VRT 10-FEB-2003
 LOCUS
 DEFINITION clone MGC:55553 IMAGE:2642773, mRNA, complete cds.
 ACCESSION BC045381
 VERSION BC045381.1 GI:28278875
 KEYWORDS
 SOURCE MGC.
 ORGANISM Dario rerio (zebrafish)
 Dario rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.

REFERENCE

1 (bases 1 to 1719)

AUTHORS
TITLE
JOURNAL

Strausberg, R.
Direct Submission
Submitted (31-Jan-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami
cDNA Library Preparation: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadane@systembiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whitting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAX Plate: 101 Row: m Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.

FEATURES
source

1..1719
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:55553 IMAGE:2642773"
/tissue_type="Whole body, adult, (one male and one female,
including unfertilized eggs)"
/clone_lib="Sugano_Kawakami_zebrafish_DNA"
/lab_host="DH10B"
/note="Vector: pME18S-FL3"
122..1270
/codon_start=1
/product="Similar to interleukin-1 receptor-associated
kinase 4"
/protein_id="AAH45381.1"
/db_xref="GI:2827876"
/translation="MSDVTPTPYRKLRYSLRALDLDPDTRWSIMADISPCGE
PRYTOMRRERFEACVLQKSPTELLFDWGTSDCTVDLVEILRHQLFAVTVLLPD
HSVCHAHTRFCEAEAPASAVLQACEITQYVADNSNKPOKISKPEVDSNAPVQEO
ELFPEVDSNAPVQEOELFPEVDSNAPVQEOELFPEVDSNAPVQEOELFPEVDSNAPVQEO
FRGMEGRKVAVKLANPLDSSYEDLKKQNOELQTLKLSHEVNLAVLCSGSPPL
CVFPELVNSLLEKLACAEHTPLATVNRNCMTVGAAGLSTLHTAHTRVYSAN
ILDEGVAKISDFGLTRSAAGAGFVWASSGSGSCVAGDEG"

CDS

BASE COUNT 492 a 399 c 477 g 351 t
ORIGIN

Alignment Scores:

Pred. No.: 3.5e-29 Length: 1719
Score: 287.00 Matches: 51
Percent Similarity: 73.68% Conservative: 19
Best Local Similarity: 53.68% Mismatches: 25
Query Match: 55.51% Indels: 0
DB: 5 Gaps: 0

US-10-001-254-6 (1-98) x BC045381 (1-1719)

Oy 1 Th-TyTtAlArGyCtLeuAanValGlyLeuileArgtYstLeuSerAappheilleaSPPro 20
Db 143 ACACCGGTTGAAAGACTCCGGTACTCCGGCTTAAGACCTTGAAGACCTGCTGGACCCG 202
Oy 21 GlnGluGlyTTPtYstYsLeuAlaValAlaIleYstYsProSerGlyYsAaSPArGTYr 40
Db 203 CAGAGACACCTGAGACATCATGCGGACATCAGACACCTCGCGGGAACCAATAC 262
Oy 41 AsnGlnpHehtisileAaGArGpHeGluAlaLeuLeuGlnThrGlyYsSerProThrSer 60
Db 263 ACAcAGATGACATGAGCGGCTTGAAGACATGCTCTGACAGGGAAGAGCTCCACCATG 322

Oy 61 GluLeuLeuPheAaPTTPGlyThrThraScyTherValGlyAaPLeuValaSPLeuLeu 80
Db 323 GAGTGTCTGTTCCACTCGGAACTCAGACTGCACCTTGAGATCTGCTGGAGATCTCTG 382
Oy 81 lGlnGlnGluPheAaPLeuAlaProAlaSerLeuLeuLeuProAaP 95
Db 383 ATCCGATCAGCTGTTCCGCGCGCTGACAGTCTGCTGCCCAT 427

RESULT 15

LOCUS

AY186092 33033 bp DNA linear PRI 02-DEC-2002
DEFINITION Homo sapiens interleukin-1 receptor-associated kinase 4 (IRAK4)

AY186092 gene, complete cds.

ACCESSION AY186092.1 GI:26000792

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

repeat_region

repeat_region

variation

variation

repeat_region

variation

variation

repeat_region

variation

variation

repeat_region

variation

variation

repeat_region

variation

variation

repeat_region

variation

variation

variation

variation

variation

variation

variation

variation

variation

variation

variation

variation

variation

variation

variation

variation

```

variation      /frequency="0.01"
                /replace="c"
2365          /gene="IRAK4"
                /frequency="0.01"
                /replace="c"
2456          /gene="IRAK4"
                /frequency="0.01"
                /replace="c"
2506          /gene="IRAK4"
                /replace="c"
repeat_region  /rpt_family="Alu"
                /rpt_type=dispersed
2940          /gene="IRAK4"
                /frequency="0.04"
                /replace="c"
3023          /gene="IRAK4"
                /frequency="0.09"
                /replace="c"
3025          /rpt_family="L2"
repeat_region  /rpt_type=dispersed
3038          /gene="IRAK4"
                /frequency="0.01"
                /replace="c"
3115          /gene="IRAK4"
                /frequency="0.01"
                /replace="c"
3129          /gene="IRAK4"
                /frequency="0.72"
                /replace="g"
3189          /gene="IRAK4"
                /frequency="0.01"
                /replace="c"
3536          /gene="IRAK4"
                /frequency="0.14"
                /replace="c"
3573          /gene="IRAK4"
                /frequency="0.01"
                /replace="t"
3819          /gene="IRAK4"
                /frequency="0.17"
                /replace="t"
3877          /gene="IRAK4"
                /frequency="0.01"
                /replace="g"
4144          /gene="IRAK4"
                /frequency="0.04"
                /replace="a"
4438          /rpt_family="MIR"
repeat_region  /rpt_type=dispersed
4602          /gene="IRAK4"
                /frequency="0.01"
                /replace="c"
4797          /gene="IRAK4"
                /frequency="0.02"
                /replace="a"
4972          /gene="IRAK4"
                /frequency="0.31"
                /replace="a"
5759          /rpt_family="L2"
repeat_region

```

```

variation      /rpt_type=dispersed
5821          /gene="IRAK4"
                /frequency="0.02"
                /replace="g"
5880          /gene="IRAK4"
                /frequency="0.01"
                /replace="c"
5954          /gene="IRAK4"
                /frequency="0.01"
                /replace="c"
5964          /rpt_family="MER1_type"
repeat_region  /rpt_type=dispersed
6023          /rpt_family="MER1_type"
                /rpt_type=dispersed
6204          /gene="IRAK4"
                /frequency="0.04"
                /replace="c"
6661          /gene="IRAK4"
                /frequency="0.04"
                /replace="a"
6998          /gene="IRAK4"
                /frequency="0.04"
                /replace="g"
7015          /rpt_family="Mariner"
repeat_region  /rpt_type=dispersed
7074          /rpt_family="Alu"
                /rpt_type=dispersed
7122          /gene="IRAK4"
                /frequency="0.01"
                /replace="a"
7147          /gene="IRAK4"
                /frequency="0.01"
                /replace="a"
7234          /gene="IRAK4"
                /frequency="0.04"
                /replace="c"
7367          /rpt_family="Mariner"
repeat_region  /rpt_type=dispersed
7392          /gene="IRAK4"
                /frequency="0.01"
                /replace="c"
7597          /gene="IRAK4"
                /frequency="0.01"
                /replace="c"
7794          /gene="IRAK4"
                /frequency="0.09"
                /replace="g"

```

Alignment Scores:

Pred. No.:	2,256-23	Length:	33033
Score:	256.00	Matches:	49
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	98.00%	Mismatches:	0
Query Match:	49.52%	Indels:	0
DB:	9	Gaps:	0

US-10-001-254-6 (1-98) x AY186092 (1-33033)

```

QY      45  ILeArgATpneGluAlaLeuLeuGlnThrGlyLysSerProThrSerGluLeuPhe 64
      :::|||||
Db      14152 TTAAGAGATTGAAAGCATTTACTCAAACTGGAAAAAGTCCCACTTCTGAATTACTGTTT 14211

QY      65  AspTrpGlyThrThrAsnCysThrValGlyAspLeuValAspLeuLeuIleGlnAsnGlu 84
      |||||
Db      14212 GACTGGGGCACCAAAATTGCACAGTTGGTGTATCTTGGATCTTTGATCCAAATGAA 14271

QY      85  PhePheAlaProAlaSerLeuLeuPro 94
      |||||
Db      14272 TTTTGTCTCTGCGAGCTTTTGTCTCCA 14301
    
```

Search completed: January 18, 2004, 04:26:13
 Job time : 1667.38 secs

BEST AVAILABLE COPY

1:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1980.DAT.*
2:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1981.DAT.*
3:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1982.DAT.*
4:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1983.DAT.*
5:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1984.DAT.*
6:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1985.DAT.*
7:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1986.DAT.*
8:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1987.DAT.*
9:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1988.DAT.*
10:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1989.DAT.*
11:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1990.DAT.*
12:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1991.DAT.*
13:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1992.DAT.*
14:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1993.DAT.*
15:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1994.DAT.*
16:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1995.DAT.*
17:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1996.DAT.*
18:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1997.DAT.*
19:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1998.DAT.*
20:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA1999.DAT.*
21:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA2000.DAT.*
22:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA2001.DAT.*
23:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA2002.DAT.*
24:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA2003.DAT.*
25:	/SIDSI/gcgdata/genseq/genseq-n-emb1/NA2004.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed

Result	No.	Score	Query	Match	Length	DB	ID	Description
	1	517	100.0	294	24	AAD40074	Human IRAK4 DD (d	
	2	517	100.0	833	21	AAA09319	Human cancer asso	
	3	517	100.0	1383	24	AAD40079	Human IRAK4 gene #	
	4	517	100.0	1668	23	AA576805	DNA encoding novel	
	5	517	100.0	2817	24	AAD40085	Human IRAK4 gene #	
	6	513	99.2	1383	22	AAAD1097	Human interleukin	
	7	476	92.1	1542	22	AAAD10198	Mouse interleukin	
	8	413	79.9	501	24	AA561608	Lung small cell ca	
	9	256	49.5	405	23	AA576803	DNA encoding novel	
	10	247	47.8	211	24	AAD40084	Human IRAK4 short	
	11	247	47.8	415	23	AA576802	DNA encoding novel	
	12	247	47.8	2213	22	AAH13798	Human cDNA sequen	
	13	134.5	26.0	408	24	ABQ56733	Human colon cancer	
	14	96.5	18.7	1791	22	AA574695	Novel protein kinase	
	15	96.5	18.7	2277	22	AA584459	Interleukin-1 recep	
	16	96.5	18.7	2288	21	AA576602	DNA encoding an ir	
	17	96.5	18.7	2293	22	AA575800	Human polynucleot	
	18	96.5	18.7	2490	23	AA586902	DNA encoding novel	
	19	84	16.2	1494	23	AA567194	DNA encoding novel	
	20	79	15.3	1806	20	AA577599	Human IRAK-2alpha	
	21	79	15.3	3459	20	AA577600	Human IRAK-2beta c	
	22	78.5	15.2	861	22	AA137224	Human musculoskele	
	23	78.5	15.2	861	25	ABX60212	cDNA encoding nove	
	24	78.5	15.2	864	22	AA137222	Human musculoskele	
	25	78.5	15.2	864	22	AA137225	Human musculoskele	
	26	78.5	15.2	864	25	ABX60210	cDNA encoding nove	
	27	78.5	15.2	864	25	ABX60213	cDNA encoding nove	
	28	74.5	14.4	652	24	ABX53369	Human eosinophil- ϵ	
	29	74.5	14.4	774	24	ABX53720	Bacillus lichenifid	
	30	72.5	14.0	2819	24	ABX53386	DNA encoding cysti	
	31	72.5	14.0	2828	24	ABO91977	Human NF- κ B activa	
	32	72.5	14.0	2828	24	ABK52388	DNA encoding cysti	
	33	72.5	14.0	2829	21	AA575235	Human ORFX ORF790	
	34	72.5	14.0	3191	23	ABV23079	Human prostate exp	
	35	72.5	14.0	3191	23	ABV28915	Human prostate exp	
	36	72	13.9	2344	23	ABL16597	Drosophila melanoc	
	37	72	13.9	4574	23	ABL16596	Drosophila melanoc	
	38	72	13.9	534720	19	AAV30458	Rhizobium species	
	39	72	13.9	536165	19	AAV30459	Rhizobium species	
	40	71.5	13.8	527	22	AAV17589	Human breast cancer	
	41	70.5	13.6	470	22	ABAA3336	Human breast cell	
	42	70.5	13.6	470	22	AAK09039	Human brain expres	
	43	70.5	13.6	470	22	AAK27496	Human bone marrow	
	44	70.5	13.6	470	22	AAI33409	Probe #2095 used t	
	45	70.5	13.6	470	22	AAI01995	Probe #1986 used t	

ALIGNMENTS

RESULT 1

ID	standard; cDNA; 294 BP.
AAD40074	

AC AAD40074 ;

DT 22-OCT-2002 (first entry)

DE Human IRAK4 DD (death domain) cDNA

KW Human; death domain; DD; death effector domain; DED; Chlamydia infection;

KW inflammation; allergy; autoimmunity; allograft rejection; cell division;

KW immunosuppressive; gene therapy; antisense therapy; gene; ss.

OS: Homo sapiens

```

XX Key Location/Qualifiers
FH 1.294
FT CDS /tag= a
FT /product= "Human IRAK4 DD"
FT /note= "No start and stop codon"
FT /partial
XX
XX MO200240680-A2.
XX
XX 23-MAY-2002.
XX
XX 15-NOV-2001; 2001WO-US44844.
XX
XX 17-NOV-2000; 2000US-0715893.
XX
XX 29-JUN-2001; 2001US-301889P.
XX
XX (BURN-) BURHAM INST.
XX
XX Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roch W,
PI Stenmer-Liwen F;
XX MPI; 2002-500222/53.
XX
XX P-PSDB; AAE24854.
XX
XX New polypeptide comprising a death domain or death effector domain,
PT useful for discovery of drugs that suppress infection, inflammation,
PT allergy, sepsis, autoimmunity, allograft rejection and other diseases
PT
XX
XX Claim 18; Page 173-174; 209p; English.
XX
XX The invention relates to an isolated polypeptide comprising a death
XX domain (DD), death effector domain (DED) or NB-ARC domain. The invention
XX is useful for identifying a binding agent, preferably a protein or a drug
XX that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
XX domain from DAP3, IRAK4, CTDD (Chlamydia trachomatis DD protein), DED4 or
XX NIDD (NGFR-interacting Death Domain), with a candidate binding agent and
XX detecting the association of the domain and the candidate binding agent,
XX by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
XX chemical crosslinking, nuclear magnetic resonance (NMR), mass
XX spectroscopy (MS) and FPA. The invention is useful for modulating the
XX level of a cell process such as cell proliferation, cell adhesion, cell
XX stress responses, responses to microbial infection and B cell
XX immunoglobulin class switching, in particular apoptosis within a cell.
XX Antibody specifically reactive with CTDD DD of C. trachomatis, C.
XX muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the
XX CTDD DD protein is useful for detecting a Chlamydia infection. The
XX invention is useful for modulating the activity of oncogenic proteins,
XX for treating a pathology caused by the oncogenic proteins and for
XX treating bacterial infections by modulating the activity of bacterial
XX proteins. The protein and antibody specific for it are useful for
XX discovery of drugs that suppress infection, inflammation, allergy,
XX sepsis, autoimmunity, allograft rejection and other diseases. The protein
XX is useful for treating immune-based pathologies, pathologies associated
XX with cell division, inflammatory diseases such as sepsis, fibrosis,
XX arthritis, graft versus host disease. The invention is used in antisense
XX therapy and gene therapy. The present sequence is human IRAK4 DD cDNA.
XX
XX
XX Sequence 294 BP; 84 A; 55 C; 62 G; 93 T; 0 other;
XX
XX Alignment Scores:
XX
XX Pred. No.: 3,94e-64 Length: 294
XX Score: 517.00 Matches: 98
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 24 Gaps: 0
XX
XX US-10-001-254-6 (1-98) x AAD40074 (1-294)
XX
XX 1 Th-TyTValArgCysLeuAenValGlyLeuILeArgLysLeuSerAapPheILeasPro 20
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```

```

DB 1 ACATATGTCGCGCTGCCTCATGTTGGACTAATTAAGAGACCTGTCAGATTTATGATCCT 60
QY 21 GInGIuGIYTPlyLeVLeuAlaValAlaIleuLysProSerGIYAspAspArgTYr 40
DB 61 CAGAGAGAGATGAGAGAGAGATTACCTTACCTATTAAAAACATCTGGTGATCATAGATAC 120
QY 41 AenGIuPheHsIleArgArgPheGIuAlaLeuLeuGInThrGIYLyseSerProThSer 60
DB 121 AATCAGTTTCACATAGAGAGATTGACACATTAAGACATTAAGCTGAGAAAAAGTCCACTTCT 180
QY 61 GIuLeuLeuPheAspTTPGIYThrThrasnCysTrValGIYAspLeuValAspLeuLeu 80
DB 181 GAATTACTGTTTGCCTGGGCGACCAAAATGACAGTTGGAGATCTGTGATCTTTTG 240
QY 81 IIGInaNGInuPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
DB 241 ATCCAAAGATGAAATTTTGTCTCTCGAGCTTTTGTCTCCAGATGCTGTCTCC 294
XX
XX RESULT 2
XX ID AAA09319 standard; DNA; 833 BP.
XX
XX AC AAA09319;
XX
XX DT 10-AUG-2000 (first entry)
XX
XX DE Human cancer associated antigen precursor DNA, clone NY-REN-64.
XX
XX KW renal cancer; cancer associated antigen precursor; diagnosis;
XX cytostatic; ss.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
FT CDS 50..670
FT /tag= a
XX
XX MO200020587-A2.
XX
XX 13-APR-2000.
XX
XX 04-OCT-1999; 99WO-US22873.
XX
XX 05-OCT-1998; 98US-0166300.
XX
XX 05-OCT-1998; 98US-0166350.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX PA Obata Y, Gout I, Tureci O, Sahin U, Pfeundschnuh M, Scanlan MJ;
PI Stockert E, Chen Y, Old LJ, Jager E, Knuth A;
XX
XX DR MPI; 2000-303774/26.
XX
XX P-PSDB; AAY92347.
XX
XX Preventing, diagnosing and/or treating disorders associated with
XX abnormal expression of human cancer associated antigens
XX
XX Claim 57; Page 85; 121p; English.
XX
XX
XX AAA09310-20 are novel genes isolated by SEREX screening from a renal
XX cancer cell line 1973/10.4. The genes encode cancer associated antigen
XX precursors. These gene products are useful in methods for preventing,
XX diagnosing and/or treating disorders, especially cancer, associated with
XX abnormal expression of human cancer associated antigens. The method
XX comprises contacting a sample from a subject with an agent that
XX specifically binds to the nucleic acid molecule or expression product
XX (or fragment) complexed with a human leukocyte antigen (HLA) molecule
XX and determining the interaction between the agent and the nucleic acid
XX molecule or the expression product as a determination of the disorder.
XX
XX
XX Sequence 833 BP; 273 A; 155 C; 179 G; 226 T; 0 other;
XX
XX Alignment Scores:

```


Pred. No.: 1.68e-63 Length: 833
 Score: 517.00 Matches: 98
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0

US-10-001-254-6 (1-98) x AAA09319 (1-833)

QY 1 ThrTYrValArgCYsLLeuAsnValGlyLeuIleArgLysLeuSerAspPheIleAspPro 20
 Db 74 ACAATATGCGCCCTGCTCAATGTTGGACTTAATTAGGAAGCTGTCAAGATTATTAATGATCTT 133
 QY 21 GlnGlnGlyTTPlySLysLeuAlaValAlaIleLysLysProSerGlyAspAspArgTyr 40
 Db 134 CAAGAAGATGGAAGAAAGATTAGCTGTACTATTAAAAACATCTGGTGGATGATGATAC 193
 QY 41 AsnGlnPheHisIleArgArgPheGlnAlaLeuLeuGlnThrGlyLysSerProThSer 60
 Db 194 AATCAGTTTCACATAGAGAGATTGGAACATTACTTCAAACTGGMAAAAGTCCACCTTCT 253
 QY 61 GluLeuLeuPheAspTTPGlyThrThrasnCysThrValGlyAspLeuValAspLeuLeu 80
 Db 254 GAATTACTGTTTGACTGGGGCAACCAATATGCAAGTTGGATCTTGTGATCTTTTG 313
 QY 81 IleGlnAsnGlnPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
 Db 314 ATCCAAATGAAATTTTTCCTCTGCGAGTCTTTTTCCTCCAGATGCTGTTCCC 367

RESULT 3

AAD40079 ID AAD40079 standard; DNA; 1383 BP.

AC AAD40079;

DT 22-OCT-2002 (first entry)

XX Human IRAK4 gene #1.

Human: death domain; DD; death effector domain; DED; Chlamydia infection; NB-ARC domain; apoptosis; oncogenic protein; Bacterial infection; sepsis; inflammation; allergy; autoimmunity; allograft rejection; cell division; immune-based pathology; fibrosis; arthritis; graft versus host disease; immunosuppressive; gene therapy; antisense therapy; gene; de.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1383

FT /*Cag= a

FT /product= "Human IRAK4"

PN W0200240680-A2.

XX 23-MAY-2002.

XX 15-NOV-2001; 2001WO-US44844.

PR 17-NOV-2000; 2000US-0715893.

XX 29-JUN-2001; 2001US-301889P.

XX (BURN-) BURNHAM INST.

XX Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;

PI Stemmer-Liwen F;

XX WPI; 2002-500222/53.

DR P-PSDB; AAE24859.

XX New polypeptide comprising a death domain or death effector domain,

PT useful for discovery of drugs that suppress infection, inflammation,

PT allergy, sepsis, autoimmunity, allograft rejection and other diseases

XX Claim 19; Page 180-182; 209pp; English.

CC The invention relates to an isolated polypeptide comprising a death
 CC domain (DD), death effector domain (DED) or NB-ARC domain. The invention
 CC is useful for identifying a binding agent, preferably a protein or a drug
 CC that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
 CC domain from DAP3, IRAK4, CTDD (Chlamydia trachomatis DD protein), DED4 or
 CC NIDP (NGFR-interacting Death Domain), with a candidate binding agent and
 CC detecting the association of the domain and the candidate binding agent,
 CC by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
 CC chemical crosslinking, nuclear magnetic resonance (NMR), mass
 CC spectroscopy (MS) and FPA. The invention is useful for modulating the
 CC level of a cell process such as cell proliferation, cell adhesion, cell
 CC stress responses, responses to microbial infection and B cell
 CC immunoglobulin class switching, in particular apoptosis within a cell.
 CC Antibody specifically reactive with CTDD DD of C. trachomatis, C.
 CC muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the
 CC CTDD DD protein is useful for detecting a Chlamydia infection. The
 CC invention is useful for modulating the activity of oncogenic proteins,
 CC for treating a pathology caused by the oncogenic proteins and for
 CC treating bacterial infections by modulating the activity of bacterial
 CC proteins. The protein and antibody specific for it are useful for
 CC discovery of drugs that suppress infection, inflammation, allergy,
 CC sepsis, autoimmunity, allograft rejection and other diseases. The protein
 CC is useful for treating immune-based pathologies, pathologies associated
 CC with cell division, inflammatory diseases such as sepsis, fibrosis,
 CC arthritis, graft versus host disease. The invention is used in antisense
 CC therapy and gene therapy. The present sequence is human IRAK4 gene.

CC Sequence 1383 BP; 463 A; 243 C; 283 G; 394 T; 0 other;

Alignment Scores:

Pred. No.: 3.4e-63 Length: 1383
 Score: 517.00 Matches: 98
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-10-001-254-6 (1-98) x AAD40079 (1-1383)

QY 1 ThrTYrValArgCYsLLeuAsnValGlyLeuIleArgLysLeuSerAspPheIleAspPro 20
 Db 25 ACAATATGCGCCCTGCTCAATGTTGGACTTAATTAGGAAGCTGTCAAGATTATTAATGATCTT 84
 QY 21 GlnGlnGlyTTPlySLysLeuAlaValAlaIleLysLysProSerGlyAspAspArgTyr 40
 Db 85 CAAGAAGATGGAAGAAAGATTAGCTGTACTATTAAAAACATCTGGTGGATGATGATAC 144
 QY 41 AsnGlnPheHisIleArgArgPheGlnAlaLeuLeuGlnThrGlyLysSerProThSer 60
 Db 145 AATCAGTTTCACATAGAGAGATTGGAACATTACTTCAAACTGGMAAAAGTCCACCTTCT 204
 QY 61 GluLeuLeuPheAspTTPGlyThrThrasnCysThrValGlyAspLeuValAspLeuLeu 80
 Db 205 GAATTACTGTTTGACTGGGGCAACCAATATGCAAGTTGGATCTTGTGATCTTTTG 264
 QY 81 IleGlnAsnGlnPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
 Db 265 ATCCAAATGAAATTTTTCCTCTGCGAGTCTTTTTCCTCCAGATGCTGTTCCC 318

RESULT 4

AAS76805 ID AAS76805 standard; cDNA; 1668 BP.

AC AAS76805;

DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #12609.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX
XX MO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG12618.
XX
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX
XX Claim 1; SEQ ID No 12609; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAG64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1668 BP; 571 A; 289 C; 332 G; 476 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 4,41e-63 length: 1668
Score: 517.00 Matches: 98
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 23 Gaps: 0
US-10-001-254-6 (1-98) x AAS76805 (1-1668)
QY 1 ThrTyrValArgCysLeuAsnValIGlyLeuIleArgLysLeuSerAaspPheIleAaspPro 20
DB 25 ACATATAGTGGCGCTGCTCAATGTTGACCTATTTAGCAAGCTGTCAGATTATTTGATCT 84
QY 21 GInGluGlyTyrPylsYsLeuAlaValAlaIleLysLysProSerGlyAaspAparGlyTyr 40
DB 85 CAAAGAGATGGAAGAGATTAGCTGTAGCTATTAAACCATCTGCTGATGATAGTAC 144
QY 41 AserGlnPheHisIleLeaArgPheGluAlaLeuLeuGlnThrGlyLysSerProThrSer 60
DB 145 AATCAGTTTCACTAAAGAAATTGGAAGCAATTCTTCAAACTGGAAAAAGTCCCACTTCT 204

QY 61 GluLeuLeuPheAaspTyrPylsYsLeuValGlyAspLeuValAspLeuLeu 80
DB 205 GAATTAAGTGTGACGAGGACCAACAAATTCACAGTGTGATCTTGTGATCTTTTG 264
QY 81 IleGlnAsnGluPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
DB 265 ATCCAAATGAATTTTGTCTCTCGAGCTTTTGTCTCCAGATGCTTCC 318
RESULT 5
AAD40085
ID AAD40085 standard; DNA; 2817 BP.
XX
XX AAD40085;
AC
XX
XX 22-OCT-2002 (first entry)
DT
XX
XX Human IRAK4 gene #2.
DE
XX
XX Human; death domain; DD; death effector domain; DED; Chlamydia infection;
KW NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
KW inflammation; allergy; autoimmunity; allograft rejection; cell division;
KW immune-based pathology; fibrosis; arthritis; graft versus host disease;
KW immunosuppressive; gene therapy; antisense therapy; gene; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH CDS 50..1432
FT /tag= a
FT /product= "Human IRAK4"
XX
XX MO200240680-A2.
XX
XX 23-MAY-2002.
PD
XX
XX 15-NOV-2001; 2001WO-US44844.
XX
XX 17-NOV-2000; 2000US-0715893.
PR 29-JUN-2001; 2001US-301889P.
XX
XX (BURN-) BURNHAM INST.
PA
XX
XX Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W,
PI Stenmer-Jiwen F;
XX
XX WPI; 2002-500222/53.
DR P-PSDB; AAE24865.
XX
XX New polypeptide comprising a death domain or death effector domain,
PT useful for discovery of drugs that suppress infection, inflammation,
PT allergy, sepsis, autoimmunity, allograft rejection and other diseases
PT -
XX
XX Claim 19; Page 194-196; 209pp; English.
XX
XX The invention relates to an isolated polypeptide comprising a death
XX domain (DD), death effector domain (DED) or NB-ARC domain. The invention
XX is useful for identifying a binding agent, preferably a protein or a drug
XX that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
XX domain from DAP3, IRAK4, CTSD (Chlamydia trachomatis DD protein), DED4 or
XX NIDD (NGFR-interacting Death Domain), with a candidate binding agent and
XX detecting the association of the domain and the candidate binding agent,
XX by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
XX chemical crosslinking, nuclear magnetic resonance (NMR), mass
XX spectroscopy (MS) and FPA. The invention is useful for modulating the
XX level of a cell process such as cell proliferation, cell adhesion, cell
XX stress responses, responses to microbial infection and B cell
XX immunoglobulin class switching, in particular apoptosis within a cell.
XX Antibody specifically reactive with CTSD DD of C. trachomatis, C.
XX muridarum, C. pneumoniae, and C. pitraci or a nucleic acid encoding the
XX CTSD DD protein is useful for detecting a Chlamydia infection. The
XX invention is useful for modulating the activity of oncogenic proteins,
XX for treating a pathology caused by the oncogenic proteins and for

CC treating bacterial infections by modulating the activity of bacterial
 CC proteins. The protein and antibody specific for it are useful for
 CC discovery of drugs that suppress infection, inflammation, allergy,
 CC sepsis, autoimmunity, allograft rejection and other diseases. The protein
 CC is useful for treating immune-based pathologies, pathologies associated
 CC with cell division, inflammatory diseases such as sepsis, fibrosis,
 CC arthritis, graft versus host disease. The invention is used in antisense
 CC therapy and gene therapy. The present sequence is human IRAK4 gene.
 XX

SO Sequence 2817 BP; 912 A; 547 C; 586 G; 772 T; 0 other;

Alignment Scores:

Pred. No.:	9.15e-63	Length:	2817
Score:	517.00	Matches:	98
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-10-001-254-6 (1-98) x AAD40085 (1-2817)

QY 1 ThrTyrValArgCysLeuAsnValGlyLeuIleArgIysLeuSerAspPheIleAspPro 20
 |||||
 DB 74 ACATATGCGGCTCCCTCAATGTGACATAATTGGAAGCTGTGAGATTATTATGATCTT 133
 |||||
 QY 21 GlnGluGlyTrpIysLeuAlaValAlaIleIysLysProSerGlyAspAspArgTyr 40
 |||||
 DB 134 CAAGAAGATGGAAGAAAGTTAGCTGTAGCTATTAAAAACCATCTGATGATGATATAC 193
 |||||
 QY 41 AsnGlnPheHisIleArgArgPheGlnAlaLeuLeuGlnThrGlyLysSerProThrSer 60
 |||||
 DB 194 AATCAGTTTCACATTAAGAGAGATTGGAACATTAATCAACTGAAAAAGTCCACTTCT 253
 |||||
 QY 61 GluLeuLeuPheAspTrpGlyThrThrAsnGlyThrValGlyAspLeuValAspLeuLeu 80
 |||||
 DB 254 GAATTACTGTTGACCTGGGACCAACAATTGACAGTTGATGATCTTGGATCTTTGG 313
 |||||
 QY 81 IleGlnAsnGluPhePheAlaProAlaSerLeuLeuProAspAlaValPro 98
 |||||
 DB 314 ATCCAAATGATTTTTCCTCCCGCAGCTTTTTCCTCCCAATGCTCTTCCC 367
 |||||

RESULT 6
 AAD10197 standard; cDNA; 1383 BP.

AC AAD10197;

DT 24-SEP-2001 (first entry)

DE Human interleukin-1 (IL-1) receptor-associated kinase (IRAK)-4 cDNA.

XX Human, interleukin-1 receptor associated kinase-4; IRAK-4; cytosolic;
 KW IL, antibacterial; antiinflammatory; ophthalmological; vasotropic; OPD;
 KW chronic obstructive pulmonary disease; neuroprotective; chronic cough;
 KW adult respiratory disease syndrome; pulmonary fibrosis; asthma; ARDS;
 KW interstitial lung disease; allergic rhinitis; transplant rejection;
 KW autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;
 KW multiple sclerosis; diabetes; cancer; solid tumour; lymphoma; stroke;
 KW cardiovascular disease; atherosclerosis; neurodegenerative disease;
 KW sepsis; osteoarthritis; rash; osteoporosis; psoriasis; dermatitis;
 KW inflammatory bowel disease; gout; Crohn's disease; ulcerative colitis;
 KW Behcet's syndrome; ankylosing spondylitis; IL-1 receptor/Toll receptor;
 KW sarcoidosis; transgenic animal; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1383

FT /*tag= a

XX /product= "Human IRAK-4"

XX MO200151641-A1.

PD 19-JUL-2001.

XX 12-JAN-2001; 2001WO-US01171.

XX 13-JAN-2000; 2000US-0176395.

PA (TUL-1) TULARIK INC.

PI Wesche H, Li S;

DR WPI, 2001-451860/48.

DR P-PSDB; AAE05398.

PT Novel human interleukin-1 receptor associated kinase polypeptide,
 useful for identifying modulators of the polypeptide for treating gout,
 PT asthma, allergic rhinitis, multiple sclerosis and skin cancer

PS Claim 7; Fig 2; 89pp; English.

XX The present sequence is a cDNA encoding human interleukin (IL)-1 receptor
 CC associated kinase (IRAK)-4. IRAK associate with activated IL-1, IL-18
 CC and other receptors and act to transduce signals originating from the
 CC activated receptors, ultimately leading to a variety of downstream
 CC effects such as nuclear factor (NF)-kappaB activation. The IRAK-4
 CC inhibitors are useful for treating inflammatory diseases such as
 CC pulmonary diseases and diseases of the airway (e.g., adult respiratory
 CC disease syndrome (ARDS), chronic obstructive pulmonary disease (OPD),
 CC pulmonary fibrosis, interstitial lung disease, asthma, chronic cough
 CC or allergic rhinitis), transplant rejection, autoimmune diseases (e.g.,
 CC rheumatoid arthritis), systemic lupus erythematosus, multiple sclerosis
 CC (diabetes), cancer (e.g., solid tumours, skin cancer or lymphoma),
 CC cardiovascular diseases (e.g., stroke and atherosclerosis), diseases
 CC of the central nervous system (e.g., neurodegenerative disease), CD4
 CC mediated sepsis, non-CD4 mediated sepsis, osteoarthritis, osteoporosis,
 CC psoriasis, diseases of the skin (e.g., rash, contact dermatitis, atopic
 CC dermatitis), inflammatory bowel disease (e.g., Crohn's disease and
 CC ulcerative colitis), Behcet's syndrome, ankylosing spondylitis, gout,
 CC sarcoidosis and ophthalmic diseases and conditions. The inhibitors of
 CC IRAK-4 activity or expression are used to inhibit signal transduction
 CC resulting from the activation of an interleukin-1 receptor (IL-1R)/Toll
 CC receptor in a cell. They also inhibit the activation of a transcription
 CC factor that activates NFkappaB in the cell. IRAK-4 is used to create a
 CC nonhuman transgenic animal which is useful for testing the function of
 CC IRAK-4 in vivo. to generate models for the study of inflammatory
 CC disorders and conditions and for the development of potential treatments
 CC for IRAK-4-related inflammatory diseases and conditions. IRAK-4 sequences
 CC are also used in gene therapy and in antisense therapy.

XX SO Sequence 1383 BP; 463 A; 244 C; 283 G; 393 T; 0 other;

Alignment Scores:

Pred. No.:	1.26e-62	Length:	1383
Score:	513.00	Matches:	97
Percent Similarity:	98.98%	Conservative:	0
Best Local Similarity:	98.98%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	22	Gaps:	0

US-10-001-254-6 (1-98) x AAD10197 (1-1383)

QY 1 ThrTyrValArgCysLeuAsnValGlyLeuIleArgIysLeuSerAspPheIleAspPro 20
 |||||
 DB 25 ACATATGCGGCTCCCTCAATGTGACATAATTGGAAGCTGTGAGATTATTATGATCTT 84
 |||||
 QY 21 GlnGluGlyTrpIysLeuAlaValAlaIleIysLysProSerGlyAspAspArgTyr 40
 |||||
 DB 85 CAAGAAGATGGAAGAAAGTTAGCTGTAGCTATTAAAAACCATCTGATGATGATATAC 144
 |||||
 QY 41 AsnGlnPheHisIleArgArgPheGlnAlaLeuLeuGlnThrGlyLysSerProThrSer 60
 |||||
 DB 145 AATCAGTTTCACATTAAGAGAGATTGGAACATTAATCAACTGAAAAAGTCCACTTCT 204
 |||||
 QY 61 GluLeuLeuPheAspTrpGlyThrThrAsnGlyThrValGlyAspLeuValAspLeuLeu 80
 |||||

```

DB      205 GAATTAAGTGTGACGGGACACCAATGACAGCTGCGATCTTGTGATCTTTG 264
QY      81 IIEGINAENGluPheheAlProAlaSerleuLeuProAspAlaValPro 98
DB      265 ATCCAAAAGAAATTTTCTCTCGAGATCTTTGCTCCAGATGCTTTCC 318

RESULT 7
AADI0198
ID      AADI0198 standard; cDNA, 1542 BP.
XX
AC      AADI0198;
XX
XX      24-SEP-2001 (first entry)
XX
DE      Mouse interleukin-1 (IL-1) receptor-associated kinase (IRAK)-4 cDNA.
XX
XX      Mouse; interleukin-1 receptor associated kinase-4; IRAK-4; cytosolic;
XX      IL; antibacterial; antiinflammatory; ophthalmological; vasotropic; OPD;
XX      chronic obstructive pulmonary disease; neuroprotective; chronic cough;
XX      adult respiratory disease syndrome; pulmonary fibrosis; asthma; ARDS;
XX      interstitial lung disease; allergic rhinitis; transplant rejection;
XX      autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;
XX      multiple sclerosis; diabetes; cancer; solid tumour; lymphoma; stroke;
XX      cardiovascular disease; atherosclerosis; neurodegenerative disease;
XX      sepsis; osteoarthritis; rash; osteoporosis; psoriasis; dermatitis;
XX      inflammatory bowel disease; gout; Crohn's disease; ulcerative colitis;
XX      Behcet's syndrome; ankylosing spondylitis; IL-1 receptor/toll receptor;
XX      sarcoidosis; transgenic animal; ss.
XX
XX      Mus sp.
XX
FH      Key      Location/Qualifiers
FT      CDS      1..1542
                /-tag= a
                /product= "Mouse IRAK-4"
XX
XX      MO200151641-A1.
XX
PD      19-JUL-2001.
XX
PF      12-JAN-2001; 2001WO-US01171.
XX
PR      13-JAN-2000; 2000US-0176395.
XX
PA      (TULSA-) TULARIK INC.
XX
PI      Wesche H, Li S;
XX
DR      WPI; 2001-451860/48.
XX
DR      P-PsDB; AAE05399.
XX
XX
PT      Novel human interleukin-1 receptor associated kinase polypeptide,
PT      useful for identifying modulators of the polypeptide for treating gout,
PT      asthma, allergic rhinitis, multiple sclerosis and skin cancer
XX
PS      Claim 26; Fig 4; 89pp; English.
XX
XX
CC      The present sequence is a cDNA encoding mouse interleukin (IL)-1 receptor
CC      associated kinase (IRAK)-4. IRAK associate with activated IL-1, IL-18
CC      and other receptors and act to transduce signals originating from the
CC      activated receptors, ultimately leading to a variety of downstream
CC      effects such as nuclear factor (NF)-kappaB activation. The IRAK-4
CC      inhibitors are useful for treating inflammatory diseases such as
CC      pulmonary diseases and diseases of the airway (e.g., adult respiratory
CC      disease syndrome (ARDS), chronic obstructive pulmonary disease (OPD),
CC      pulmonary fibrosis, interstitial lung disease, asthma, chronic cough
CC      or allergic rhinitis), transplant rejection, autoimmune diseases (e.g.,
CC      rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis
CC      or diabetes), cancer (e.g., solid tumours, skin cancer or lymphoma),
CC      cardiovascular diseases (e.g., stroke and atherosclerosis), diseases
CC      of the central nervous system (e.g., neurodegenerative disease), CD14
CC      mediated sepsis, non-CD14 mediated sepsis, osteoarthritis, osteoporosis,

```

```

CC      psoriasis, diseases of the skin (e.g., rash, contact dermatitis, atopic
CC      dermatitis), inflammatory bowel disease (e.g., Crohn's disease and
CC      ulcerative colitis), Behcet's syndrome, ankylosing spondylitis, gout,
CC      sarcoidosis and ophthalmic diseases and conditions. The inhibitors of
CC      IRAK-4 activity or expression are used to inhibit signal transduction
CC      resulting from the activation of an interleukin-1 receptor (IL-1R)/Toll
CC      receptor in a cell. They also inhibit the activation of a transcription
CC      factor that activates NFkappaB in the cell. IRAK-4 is used to create a
CC      nonhuman transgenic animal which is useful for testing the function of
CC      IRAK-4 in vivo, to generate models for the study of inflammatory
CC      disorders and conditions and for the development of potential treatments
CC      for IRAK-4-related inflammatory diseases and conditions. IRAK-4 sequences
CC      are also used in gene therapy and in antisense therapy.
XX
SQ      Sequence 1542 BP; 421 A; 392 C; 423 G; 306 T; 0 other;
XX
Alignment Scores:
Pred. No.:      2.64e-57      Length:      1542
Score:          476.00      Matches:      89
Percent Similarity: 95.92%      Conservative: 5
Best Local Similarity: 90.82%      Mismatches: 4
Query Match:    92.07%      Indels:      0
DB:             22      Gaps:      0

US-10-001-254-6 (1-98) x AADI0198 (1-1542)
QY      1 ThrTYValArgCYLeuAsnValGlyLeuIleArgLYLeuSerAspPheIleAspPro 20
DB      187 ACNTACATRCGACACCTTAATGTGGGATCCTTAGAAGCTGTCCGATTTTATATCT 246
QY      21 GINGIUGLYTPLeuLYLeuAlaValAlaIleLYLeuSPROSeRGlyAspAspArgTYr 40
DB      247 CAAGAAGGCTGGAAAGAAATTAGACAGTACTCAAAAAGCCGTCGCGACACAGATAC 306
QY      41 AENGluPheHeIleArgArgPheGluAlaLeuLeuGlnThrGlyLYLeuSerProThrSer 60
DB      307 AATCAGTTCCATPACAGAGATTGACAGCTTACTTCAGACCGGAGAGGCCCACTGT 366
QY      61 GluLeuLeuPheAspTTPGlyThrThrasnCYSThrValGlyAspLeuValAspLeuLeu 80
DB      367 GAACGCTGTGTGACTGGGACCAACGAACTGACAGTTGGGAGACTTGTGATCTACTG 426
QY      81 IIEGINAENGluPheheAlProAlaSerleuLeuProAspAlaValPro 98
DB      427 GTCCAGATTGACGCTGTTGCCCCCGCACTCTCTGCTGCGAGATGCCGTTCC 480

RESULT 8
AAS61608
ID      AAS61608 standard; cDNA, 501 BP.
XX
AC      AAS61608;
XX
XX
XX      29-JUN-2002 (first entry)
XX
DE      Lung small cell carcinoma antigen, cDNA #149.
XX
XX
XX      Human; cytosolic; antitumour; lung small cell cancer antigen;
XX      tumour; lung cancer; ss.
XX
XX
OS      Homo sapiens.
XX
PN      MO200177168-A2.
XX
PD      18-OCT-2001.
XX
PF      11-APR-2001; 2001WO-US11859.
XX
PR      11-APR-2000; 2000US-196780P.
XX
PR      21-JUN-2000; 2000US-213361P.
XX
PR      01-SEP-2000; 2000US-229763P.
XX
PR      05-SEP-2000; 2000US-230629P.
XX
PR      14-SEP-2000; 2000US-232565P.
XX
PR      19-DEC-2000; 2000US-257037P.

```

PR 08-JAN-2001; 2001US-260796P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Lodes MJ, Wang T, Mohamath R, Indirias CY;
PI WPI; 2002-010896/01.
XX
XX Lung tumour polynucleotide and polypeptides useful in therapy and
PT diagnosis of cancer especially lung cancer
XX
XX Claim 1; Page 174; 295pp; English.
XX
XX The invention relates to novel isolated lung small cell cancer antigen
CC polynucleotides (I) and polypeptides (II) used in a method of detecting
CC cancer in a patient. The method is optionally performed by
CC utilizing oligonucleotides (III) where the biological sample
CC from the patient is contacted with (III), detecting the amount of
CC polynucleotide hybridised to (III) in the sample and comparing the
CC amount of polynucleotide to a predetermined cut-off value and thereby
CC determining cancer in a patient. (I), (II) or antigen-presenting cells
CC expressing (II) is useful for stimulating and/or expanding T cells
CC specific for a tumour protein. The method comprises contacting T cells
CC with one of the components under conditions to permit the stimulation
CC and/or expansion of the cells. A composition comprising (I) is useful for
CC stimulating an immune response in a patient and for inhibiting the
CC development of a cancer especially lung cancer in a patient. An
CC isolated T cell population is useful for removing tumour cells from the
CC biological sample and for inhibiting the development of cancer in a
CC patient. AAS61460-AAS61874 represent novel human lung small cell
CC cancer antigen coding sequences of the invention.
XX
SQ Sequence 501 BP; 166 A; 102 C; 97 G; 134 T; 2 other;
Alignment Scores:
Pred. No.: 4.94e-49 Length: 501
Score: 413.00 Matches: 93
Percent Similarity: 94.00% Conservative: 1
Best Local Similarity: 93.00% Mismatches: 4
Query Match: 79.88% Indels: 4
DB: Gaps: 0
US-10-001-254-6 (1-98) x AAS61608 (1-501)
QY 1 ThrtYrValaArGcYsLeuAanValaGlyLeuIleArGlyLeuSerAaspPheIleAaspPro 20
DB 45 ACATATGTGCGCTCCCTCAAGTGGAGCTAATTAAGAAAGCTGACATTTATTGATCCT 104
QY 21 GlnGluGlyTrrPlySlyLeuAlaValAlaIleLysLysProSerGlyAaspAparGlyTyr 40
DB 105 CAAGAAAGATGGAAGAAAGTTAGCTGTAGCTATTAAAAAACATCTGGTGAATGATGATAC 164
QY 41 AenGln-PheHisIleArGargPheGluAlaLeuLeuGlnThrGlyLysSerProThrSe 60
DB 165 AATCAAGCTTTCACATTAAGAGATTGAAGCATT-CTTCAAACTGAAAAAGATCCCACTTC 223
QY 60 rGluLeuLeuPheAspTrp-GlyThrThrAsnCystrValaGlyAaspLeuValaAspLeu 80
DB 224 TTGAATATCTGTTTACGCGGCGCACCAAAATTGCAAGTTGGATCTTGCGATCTTT 283
QY 80 euIleGlnAenGluPhePheAlaProAlaSerLeuLeuLeuProAaspAlaValPro 98
DB 284 TGAATCCAAATGAATT-TTTCGCTCTCGAGTCTTTTGTCTCCCAAGTCGTCTCC 338
RESULT 9
AAS76803
ID AAS76803 standard; cDNA; 405 BP.
XX
XX AAS76803;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #12607.

XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
DR P-PSDB; ABG12616.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 1; SEQ ID No 12607; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 405 BP; 93 A; 87 C; 83 G; 142 T; 0 other;
Alignment Scores:
Pred. No.: 7.44e-27 Length: 405
Score: 256.00 Matches: 49
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.00% Mismatches: 0
Query Match: 49.52% Indels: 0
DB: Gaps: 0
US-10-001-254-6 (1-98) x AAS76803 (1-405)
QY 45 IleArgArgPheGluAlaLeuLeuGlnThrGlyLysSerProThrSerGluLeuPhe 64
DB 114 TTAAAGAGATTGAAACATTAATTAAGTCAAAAGTCCCACTTGAATTAATGATT 173
QY 65 AaspTrpGlyThrThrAsnCystrValaGlyAaspLeuValaAspLeuLeuIleGlnAenGlu 84
DB 174 GACTGGGGGACCAACAATTGCAAGTTGGATCTTGCGATCTTTGATCCAAATGA 233
QY 85 PhePheAlaProAlaSerLeuLeuPro 94

DB 234 TTTTGTCTCGCGAGCTCTTTGTCCCA 263

RESULT 10

ID AAD40084 standard; DNA; 211 BP.

XX AAD40084;

XX 22-OCT-2002 (first entry)

DE Human IRAK4 short gene.

XX Human; death domain; DD; death effector domain; DED; Chlamydia infection;

KW NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;

KW inflammation; allergy; autoimmunity; allograft rejection; cell division;

KW immune-based pathology; fibrosis; arthritis; graft versus host disease;

KW immunosuppressive; gene therapy; antisense therapy; gene; ds.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..180

FT /tag= a

FT /product= "Human IRAK4 short protein"

PN MO200240680-A2.

PD 23-MAY-2002.

XX 15-NOV-2001; 2001WO-US44844.

XX 17-NOV-2000; 2000US-0715893.

PR 29-JUN-2001; 2001US-301889P.

XX (BURN-) BURNHAM INST.

PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W,

PI Stenener-Jiwen F;

XX MPI; 2002-500222/53.

DR P-PSDB; AAE24864.

XX New polypeptide comprising a death domain or death effector domain,

PT useful for discovery of drugs that suppress infection, inflammation,

PT allergy, sepsis, autoimmunity, allograft rejection and other diseases

PT -

XX Claim 19; Page 195; 209pp; English.

XX The invention relates to an isolated polypeptide comprising a death

CC domain (DD), death effector domain (DED) or NB-ARC domain. The invention

CC is useful for identifying a binding agent, preferably a protein or a drug

CC that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC

CC domain from DAP3, IRAK4, CTDD (Chlamydia trachomatis DD protein), DED4 or

CC NIDD (NGPR-interacting Death Domain), with a candidate binding agent and

CC detecting the association of the domain and the candidate binding agent,

CC by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or

CC chemical crosslinking, nuclear magnetic resonance (NMR), mass

CC spectroscopy (MS) and FPA. The invention is useful for modulating the

CC level of a cell process such as cell proliferation, cell adhesion, cell

CC stress responses, responses to microbial infection and B cell

CC immunoglobulin class switching, in particular apoptosis within a cell.

CC Antibody specifically reactive with CTDD DD of C. trachomatis, C.

CC muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the

CC CTDD DD protein is useful for detecting a Chlamydia infection. The

CC invention is useful for modulating the activity of oncogenic proteins,

CC for treating a pathology caused by the oncogenic proteins and for

CC treating bacterial infections by modulating the activity of bacterial

CC proteins. The protein and antibody specific for it are useful for

CC discovery of drugs that suppress infection, inflammation, allergy,

CC sepsis, autoimmunity, allograft rejection and other diseases. The protein

CC is useful for treating immune-based pathologies, pathologies associated

CC with cell division, inflammatory diseases such as sepsis, fibrosis,

CC arthritis, graft versus host disease. The invention is used in antisense

CC therapy and gene therapy. The present sequence is human IRAK4 short gene.

XX

SQ Sequence 211 BP; 77 A; 41 C; 36 G; 57 T; 0 other;

Alignment Scores:

Pred. No.:	5,7e-26	Length:	211
Score:	247.00	Matches:	46
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	47.78%	Indels:	0
DB:	24	Gaps:	0

US-10-001-254-6 (1-98) x AAD40084 (1-211)

QY 1 TTTTGTCTCGCGAGCTCTTTGTCCCA 263

DB 25 ACATATGTGCGCTGCTCAATGTGACTTAATTAGCAAGCTGTGAGATTATTAATCT 84

QY 21 GINGIUGLYTTPYLYSLYLEUALAVAlAlletlyslsProSerGlyAspAspArgTYr 40

DB 85 CAAGAAGATGGAAGAAGTTAGCTTACTTNTAAAAAACATCTGGTGATGATGATTC 144

QY 41 AaGinPheHisIleArg 46

DB 145 AATCAGTTTCACATAGA 162

RESULT 11

ID AAS76802 standard; CDNA; 415 BP.

XX AAS76802;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #12606.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS MO200175067-A2.

PN 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

PI MPI; 2001-639362/73.

DR P-PSDB; ABG12615.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

XX Claim 1; SEQ ID No 12606; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (ii) and its binding partners are useful in medical
CC imaging of sites expressing (ii). (i) and (ii) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 415 BP; 144 A; 83 C; 78 G; 110 T; 0 other;

Alignment Scores:

Pred. No.:	1,46e-25	Length:	415
Score:	247.00	Matches:	46
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	47.78%	Indels:	0
DB:	23	Gaps:	0

US-10-001-254-6 (1-98) x AAS76802 (1-415)

QY 1 ThrTyrValArgCysLeuAsnValGlyLeuIleArgLysLeuSerAspPheIleAspPro 20

DB 60 ACATATGTGGCCTCCCAATGTTGACTAATAGAAAGCTGACAGATTATTCATCT 119

QY 21 GlnGluGlyTrrPlyLysLeuAlaValAlaIleLysLysProSerGlyAspAspArgTyr 40

DB 120 CAAAGAAGATGGAAGAAGTAGCTGTAGCTATTAAAAACCATCTGGTGAATAGATAC 179

QY 41 AsnGlnPheHisIleArg 46

DB 180 AATCAGTTTCACATPAGA 197

RESULT 12

AAH13798

XX AAH13798 standard; cDNA; 2213 BP.

AC AAH13798;

XX 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:10742.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

XX EPI074617-A2.

PN 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELT-) HELIX RES INST.

XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Iishi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

XX Claim 8; SEQ ID 10742; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

XX full-length cDNAs defined in the specification. Where a primer set

XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in

XX in the specification. The primer sets can be used in antisense therapy and

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

XX the full-length cDNAs. The primers allow obtaining of the full-length

XX cDNAs easily without any specialized methods. AAH03166 to AAH13628 and

XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

XX represent oligonucleotides, all of which are used in the exemplification

XX of the present invention.

SQ Sequence 2213 BP; 642 A; 488 C; 516 G; 567 T; 0 other;

Alignment Scores:

Pred. No.:	1.5e-24	Length:	2213
Score:	247.00	Matches:	46
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	47.78%	Indels:	0
DB:	22	Gaps:	0

US-10-001-254-6 (1-98) x AAH13798 (1-2213)

QY 1 ThrTyrValArgCysLeuAsnValGlyLeuIleArgLysLeuSerAspPheIleAspPro 20

DB 775 ACATATGTGGCCTCCCAATGTTGACTAATAGAAAGCTGACAGATTATTCATCT 834

QY 21 GlnGluGlyTrrPlyLysLeuAlaValAlaIleLysLysProSerGlyAspAspArgTyr 40

DB 835 CAAAGAAGATGGAAGAAGTAGCTGTAGCTATTAAAAACCATCTGGTGAATAGATAC 894

QY 41 AsnGlnPheHisIleArg 46

DB 895 AATCAGTTTCACATPAGA 912

RESULT 13

ABQ56733

XX ABQ56733 standard; cDNA; 408 BP.

XX 02-AUG-2002 (first entry)

XX Human colon cancer related nucleotide sequence SEQ ID NO:428.

XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;

XX genetic analysis; diagnostic; antisense therapy; gene; ss.

XX Homo sapiens.

XX WO200229086-A2.

XX 11-APR-2002.

XX 02-OCT-2001; 2001WO-US30732.

XX

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 18, 2004, 04:27:15 ; Search time 146.603 Seconds
(without alignments)
2356.207 Million cell updates/sec

Title: US-10-001-254-6
Perfect score: 517
Sequence: 1 TYVCLNVGLIRKLSDFIDP.....LLIQNEFPAPASLLLPDAVP 98

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2324096 segs, 1762381658 residues
Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cg2_1/USP10.spool/US10001254/runcat_16012004_152426_19815/app_query.fasta_1.1109
-DB=Published Applications NA -QFMT=fastap -SUFFIX=tmpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blomsu62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MATELEN=2000000000 -USER=US10001254 @CGN_1.1.534 @runcat_16012004_152426_19815
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cg2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cg2_6/ptodata/1/pubpna/PTC_NEW_PUB.seq:*
3: /cg2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cg2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cg2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cg2_6/ptodata/1/pubpna/PTCUS_PUBCOMB.seq:*
7: /cg2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cg2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cg2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cg2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
11: /cg2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cg2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cg2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cg2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cg2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cg2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cg2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cg2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	517	100.0	294	15	US-10-001-254-5	Sequence 5, Appl1
2	517	100.0	1383	15	US-10-001-254-15	Sequence 15, Appl1
3	517	100.0	2817	11	US-09-966-451-3	Sequence 3, Appl1
4	517	100.0	2817	15	US-10-001-254-27	Sequence 27, Appl1
5	513	99.2	1383	11	US-09-759-595-2	Sequence 2, Appl1
6	476	92.1	1542	11	US-09-759-595-4	Sequence 4, Appl1
7	413	79.9	501	9	US-09-833-790-149	Sequence 149, Appl
8	256	49.5	31000	11	US-09-966-451-10	Sequence 10, Appl
9	247	47.8	211	15	US-10-001-254-25	Sequence 25, Appl
10	96.5	18.7	2288	9	US-09-863-549-1	Sequence 1, Appl1
11	96.5	18.7	2293	14	US-10-098-841-3	Sequence 3, Appl1
12	95	15.4	1888	13	US-10-340-545-1	Sequence 1, Appl1
13	79.5	15.2	864	15	US-10-198-846-1432	Sequence 1432, Ap
14	79	15.3	1782	13	US-10-366-288-19	Sequence 19, Appl
15	79	15.3	1806	9	US-09-773-753-1	Sequence 1, Appl1
16	79	15.3	3459	9	US-09-773-753-3	Sequence 3, Appl1
17	78.5	15.2	479	9	US-09-773-753-14	Sequence 14, Appl
18	78.5	15.2	861	10	US-09-764-877-3589	Sequence 3589, Ap
19	78.5	15.2	864	10	US-09-764-877-3587	Sequence 3587, Ap
20	78.5	15.2	864	10	US-09-764-877-3590	Sequence 3590, Ap
21	75.5	14.6	3128	12	US-10-108-260A-709	Sequence 709, App
22	74.5	14.4	774	10	US-09-974-300-1011	Sequence 1011, Ap
23	73	14.1	497	13	US-10-027-632-50723	Sequence 50723, A
24	73	14.1	497	13	US-10-027-632-50723	Sequence 50723, A
25	73	14.1	497	14	US-10-027-632-50723	Sequence 50723, A
26	73	14.1	497	14	US-10-027-632-50724	Sequence 50724, A
27	73	14.1	499	13	US-10-027-632-72567	Sequence 72567, A
28	73	14.1	499	13	US-10-027-632-72567	Sequence 72567, A
29	73	14.1	499	14	US-10-027-632-72567	Sequence 72567, A
30	73	14.1	499	14	US-10-027-632-72568	Sequence 72568, A
31	72.5	14.0	2828	13	US-10-024-298A-132	Sequence 132, App
32	72.5	14.0	2828	13	US-10-042-211A-132	Sequence 132, App
33	72	13.9	53615	11	US-09-939-964-1	Sequence 1, Appl1
34	71	13.7	877	13	US-10-316-253-288	Sequence 288, App
35	71	13.7	889	13	US-10-316-253-286	Sequence 286, App
36	70.5	13.6	470	9	US-09-864-761-10114	Sequence 10114, A
37	70.5	13.6	2048	12	US-10-108-260A-2016	Sequence 2016, Ap
38	70	13.5	601	13	US-10-027-632-141792	Sequence 141792, A
39	70	13.5	601	14	US-10-027-632-141792	Sequence 141792, A
40	69.5	13.4	1416	13	US-10-025-145A-19	Sequence 19, Appl
41	69.5	13.4	1785	9	US-09-887-586A-49	Sequence 49, Appl
42	69.5	13.4	1785	9	US-09-903-112-49	Sequence 49, Appl
43	69.5	13.4	1785	11	US-09-900-797-49	Sequence 49, Appl
44	69.5	13.4	1785	14	US-10-041-007-19	Sequence 19, Appl
45	69	13.3	773	15	US-10-198-846-5618	Sequence 5618, Ap

ALIGNMENTS

RESULT 1
US-10-001-254-5
Sequence 5, Application US/10001254
Publication No. US20030049702A1
GENERAL INFORMATION:
APPLICANT: Reed, John C.
APPLICANT: Godzik, Adam
APPLICANT: Pawlowski, Krzysztof
APPLICANT: Fiorentino, Loredana
APPLICANT: Lee, Sug Hyung
APPLICANT: Roth, Wilfred
APPLICANT: Steiner-Liwen, Frank
TITLE OR INVENTION: No. US20030049702A1e1 Death Domain Proteins
FILE REFERENCE: P-LT 5037
CURRENT APPLICATION NUMBER: US/10/001,254
CURRENT FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/301,889
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/715,893
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 294


```

DB      314 ATCCAAATGATTTTGTCTCTGCGAGCTTTTGTCTCCCAAGTGTCTTCCC 367
RESULT 4
US-10-001-254-27
; Sequence 27, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sug Hyung
; APPLICANT: Roth, Wilfried
; APPLICANT: Stenmer-Liwen, Frank
; TITLE OF INVENTION: No. US20030049702A1e1 Death Domain Proteins
; FILE REFERENCE: P-LJ 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 2817
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)...(1429)
US-10-001-254-27

Alignment Scores:
Pred. No.: 5.01e-71 Length: 2817
Score: 517.00 Matches: 98
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-10-001-254-6 (1-98) x US-10-001-254-27 (1-2817)
QY      1 ThrTYValaRGcysleuAnValIGlyleuIleArGlyseuSerAspPheIleAspPro 20
DB      74 ACATATGTCGCTGCTCCTCAATGTTGACATTAAGAGCTGTCACATTTTATTTGATCCT 133
QY      21 GInGIuGIYTrpLySylsleuAlaValAlaIleLySylsPProSeRGlyAspAspArgTYr 40
DB      134 CAAGAAGATGAGAAAGTACGTAGCTAGCTATTAAAAACCATCTGCTGATGATAGATAC 193
QY      41 AenGIuPheHsiIleArGArGpHeGIuAlaLeuLeuGInThrGIlyseuSerProThrSer 60
DB      194 AATCACTTTCACATTAAGAGATTGAAAGCATTAATCAACTGGAAGAAAGTCCACATCTCT 253
QY      61 GluLeuLeuPheAspTrpGIYThrThraenCySThrValAGlyAspLeuValaAspLeu 80
DB      254 GAATTAAGTGTGTACTGGGGCACCACAAATTTGACAGTTGGTATCTTGTGGATCTTTTG 313
QY      81 lIeGIaenGIuPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
DB      314 ATCCAAATGATTTTGTCTCTGCGAGCTTTTGTCTCCCAAGTGTCTTCCC 367

RESULT 5
US-09-759-595-2
; Sequence 2, Application US/09759595
; Publication No. US20030059916A1
; GENERAL INFORMATION:
; APPLICANT: Wesche, Holger
; APPLICANT: Li, Shyun
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: IRAK-4: Compositions and Methods of Use

```

```

; FILE REFERENCE: 018781-003910US
; CURRENT APPLICATION NUMBER: US/09/759,595
; CURRENT FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: US 60/176,395
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human IL-1 receptor-associated kinase 4 (IRAK-4)
; OTHER INFORMATION: CDNA
; NAME/KEY: CDS
; LOCATION: (1)..(1383)
US-09-759-595-2

Alignment Scores:
Pred. No.: 7.76e-71 Length: 1383
Score: 513.00 Matches: 97
Percent Similarity: 98.98% Conservative: 0
Best Local Similarity: 98.98% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 11 Gaps: 0

US-10-001-254-6 (1-98) x US-09-759-595-2 (1-1383)
QY      1 ThrTYValaRGcysleuAnValIGlyleuIleArGlyseuSerAspPheIleAspPro 20
DB      25 ACATATGTCGCTGCTCCTCAATGTTGACATTAAGAGCTGTCACATTTTATTTGATCCT 84
QY      21 GInGIuGIYTrpLySylsleuAlaValAlaIleLySylsPProSeRGlyAspAspArgTYr 40
DB      85 CAAGAAGATGAGAAAGTACGTAGCTAGCTATTAAAAACCATCTGCTGATGATAGATAC 144
QY      41 AenGIuPheHsiIleArGArGpHeGIuAlaLeuLeuGInThrGIlyseuSerProThrSer 60
DB      145 AATCACTTTCACATTAAGAGATTGAAAGCATTAATCAACTGGAAGAAAGTCCACATCTCT 204
QY      61 GluLeuLeuPheAspTrpGIYThrThraenCySThrValAGlyAspLeuValaAspLeu 80
DB      205 GAATTAAGTGTGTACTGGGGCACCACAAATTTGACAGTTGGTATCTTGTGGATCTTTTG 264
QY      81 lIeGIaenGIuPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
DB      265 ATCCAAATGATTTTGTCTCTGCGAGCTTTTGTCTCCCAAGTGTCTTCCC 318

RESULT 6
US-09-759-595-4
; Sequence 4, Application US/09759595
; Publication No. US20030059916A1
; GENERAL INFORMATION:
; APPLICANT: Wesche, Holger
; APPLICANT: Li, Shyun
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: IRAK-4: Compositions and Methods of Use
; FILE REFERENCE: 018781-003910US
; CURRENT APPLICATION NUMBER: US/09/759,595
; CURRENT FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: US 60/176,395
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: murine IL-1 receptor-associated kinase 4 (IRAK-4)
; OTHER INFORMATION: CDNA

```

FEATURE:
 NAME/KEY: CDS
 LOCATION: (163) (1542)
 OTHER INFORMATION: murine IRAK-4
 US-09-759-595-4

Alignment Scores:
 Pred. No.: 6,766-65 Length: 1542
 Score: 476.00 Matches: 89
 Percent Similarity: 95.92% Conservative: 5
 Best Local Similarity: 90.82% Mismatches: 4
 Query Match: 92.07% Indels: 0
 DB: 11 Gaps: 0

US-10-001-254-6 (1-98) x US-09-759-595-4 (1-1542)

QY 1 ThrTYrValArgCYsLeuAenValGlyLeuIleArgLYsLeuSerAspPheIleAspPro 20
 DB 187 ACATACATACGACACCTTAATGCGGAGATCCTTAGAGAGCTGCGATTATTGATCCT 246
 QY 21 GInGluGlyTPRlySLySLyLeuAlaValAlaIleLYsLYsProSerGlyAspAspArgTYr 40
 DB 247 CAGAAGAGGTGGAGAAATTAGAGTAGCTATCAAAAGCCGTCCGGCAGCAGCAGATAC 306
 QY 41 AenGlnPheHisIleArgArgPheGluAlaLeuGlnThrGlyLYsSerProThrSer 60
 DB 307 AATCAGTTCCATATAAGAGATTCCAGAGCTTACTTCAGACCGGAGAGAGCCCACTGT 366
 QY 61 GLeuLeuPheAspTPRGLyThrThrAsnCYsThrValGlyAspLeuValAspLeuLeu 80
 DB 367 GAACGCTGCTTGGACTGGGGCAGCAGAACTGCACAGTGGGACCTGTGTGATCTACTG 426
 QY 81 IleGlnAsnGluPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
 DB 427 GTCCGATTGACCTGTTTCCGCCCTCCTCCTGCTGCTCCGAGTGCCTGCC 480

RESULT 7

US-09-833-790-149
 Sequence 149, Application US/09833790
 Patent No. US20020068288A1
 GENERAL INFORMATION:
 APPLICANT: Lodes, Michael J.
 APPLICANT: Wang, Tonglong
 APPLICANT: Secrist, Heather
 APPLICANT: Mohamath, Raodoh
 APPLICANT: Indrias, Carol Y.
 APPLICANT: Fan, Liqun
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 FILE REFERENCE: 210121.512
 CURRENT APPLICATION NUMBER: US/09/833,790
 CURRENT FILING DATE: 2001-04-11
 NUMBER OF SEQ ID NOS: 440
 SOFTWARE: FastSeq for windows Version 4.0
 SEQ ID NO 149
 LENGTH: 501
 TYPE: DNA
 ORGANISM: Homo sapien
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1) ... (501)
 OTHER INFORMATION: n = A,T,C or G
 US-09-833-790-149

Alignment Scores:

Pred. No.: 1,336-55 Length: 501
 Score: 413.00 Matches: 93
 Percent Similarity: 94.00% Conservative: 1
 Best Local Similarity: 93.00% Mismatches: 4
 Query Match: 79.88% Indels: 0
 DB: 9 Gaps: 0

US-10-001-254-6 (1-98) x US-09-833-790-149 (1-501)

QY 1 ThrTYrValArgCYsLeuAenValGlyLeuIleArgLYsLeuSerAspPheIleAspPro 20
 DB 45 ACATATGCGCCTCCCTCAATGTTGACTAATTAGGAAGCTGCAGATTTTATGATCCT 104
 QY 21 GInGluGlyTPRlySLySLyLeuAlaValAlaIleLYsLYsProSerGlyAspAspArgTYr 40
 DB 105 CAGAAGAGATGAGAAATTAGCTGTAGCTATTAATAAACCATCTGTGTAGATAGATAC 164
 QY 41 AenGln-PheHisIleArgArgPheGluAlaLeuGlnThrGlyLYsSerProThrSe 60
 DB 165 AATCAGTTTACATTAAGAGATTGAAAGCATTTCTTCAAACTGAAAAGTCCCACTTC 223
 QY 60 rGluLeuLeuPheAspTPR-GlyThrThrAsnCYsThrValGlyAspLeuValAspLeuL 80
 DB 224 TTGAATACTGTTTACCTGGGCGACCAAAATTGACAGTTGGATCTTGTGATCTTT 283
 QY 80 eullleGlnAsnGluPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
 DB 284 TGATCCAAAATGAATT-TTGTCTCTCGAGTCTTTGCTCCAGATGCTGTCCC 338

RESULT 8

US-09-966-451-10
 Sequence 10, Application US/09966451
 Publication No. US20030087856A1
 GENERAL INFORMATION:
 APPLICANT: C. Frank Bennett
 APPLICANT: Susan M. Freiler
 TITLE OF INVENTION: ANTISENSE MODULATION OF IL-1 RECEPTOR-ASSOCIATED KINASE-4 EXPRESS
 FILE REFERENCE: RFS-0324
 CURRENT APPLICATION NUMBER: US/09/966,451
 CURRENT FILING DATE: 2001-09-28
 NUMBER OF SEQ ID NOS: 88
 SEQ ID NO 10
 LENGTH: 31000
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 US-09-966-451-10

Alignment Scores:
 Pred. No.: 4,26-28 Length: 31000
 Score: 256.00 Matches: 49
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.00% Mismatches: 0
 Query Match: 49.52% Indels: 0
 DB: 11 Gaps: 0

US-10-001-254-6 (1-98) x US-09-966-451-10 (1-31000)

QY 45 IleArgArgPheGluAlaLeuGlnThrGlyLYsSerProThrSerGluLeuPhe 64
 DB 12963 TTAAGAGATTGGAAGCATTTCTTCAACTGGAAGATCCACTTCTGAATTACTGTTT 13022
 QY 65 AspTPRGLyThrThrAsnCYsThrValGlyAspLeuValAspLeuLeuIleGlnAsnGlu 84
 DB 13023 GACTGGGCGACCACAAAATTGACAGTGTGATCTTGTGATCTTTGATCCAAATGAA 13082
 QY 85 PhePheAlaProAlaSerLeuLeuLeuPro 94
 DB 13083 TTTTGTCTCTCGAGATCTTTGTCTCCA 13112

RESULT 9

US-10-001-254-25
 Sequence 25, Application US/10001254
 Publication No. US20030049702A1
 GENERAL INFORMATION:
 APPLICANT: Reed, John C.
 APPLICANT: Godzik, Adam
 APPLICANT: Pawlowski, Krzysztof
 APPLICANT: Fiorentino, Loredana
 APPLICANT: Lee, Sug Hyung
 APPLICANT: Roth, Wilfred

Pred. No.:	0.000202	Length:	2288
Pragmatic Scores:		Matches:	23
Score:	96.50	Conservative:	12
Percent Similarity:	47.95%	Mismatches:	31
Best Local Similarity:	31.5%	Indels:	7
Query Match:	18.67%	Gaps:	1
DB:	9		

US-10-001-254-6 (1-98) X US-10-098-841-3 (1-2293)

QY 8 VALGILYLeuILEaRGLysLeuSERASPheILEaSPROGLINGLYTYRPLYSLeu 27

```

Db      88 CTCGAGAGCTGCGCTGTTCTTGACAGCTGCAGCGCGCTGGGCTGCGCGGCTG 147
Qy      28 A1aValAla1lelylsyrsProserGlyAspAspArgTyraSnglnPheHis1leArg 47
Db      148 GCAGAGAGACTTTCAGC-----AGCTGGCTGATGTTCTGCTCAT 186
Qy      48 PheGluAla1leuSnglnThrglylsyrsProThrsSerGlnleuSnglnPheAspTyrGly 67
Db      187 ATTGAAAGATGATCTAACCAGGTAAAGTGAACAGAGAAATTAATCTTGCTGCTGCGCA 246
Qy      68 ThrtThrsnCysThrsValGlyAspLeuValAspLeuLeu 80
Db      247 CAGAAAAACAAGCACCATCGGTGACCTTTTAAACAGTCTTC 285

RESULT 12
US-10-340-545-1
; Sequence 1, Application US/10340545
; Publication No. US20030157539A1
; GENERAL INFORMATION:
; APPLICANT: Flavell et al.
; TITLE OF INVENTION: IRAK-M IS A NEGATIVE REGULATOR OF TOLL-LIKE RECEPTOR SIGNALING
; FILE REFERENCE: YU-P01-011
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: 60/348176
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1888
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-340-545-1

Alignment Scores:
Pred. No.: 0.000265 Length: 1888
Score: 95.00 Matches: 24
Percent Similarity: 48.65% Conservative: 12
Best Local Similarity: 32.43% Mismatches: 28
Query Match: 18.38% Indels: 10
DB: 13 Gaps: 2

US-10-001-254-6 (1-98) x US-10-340-545-1 (1-1888)
Qy      10 Leu1leArglylsyrsEsrAspPhe1leAspProGlnGlu-----GlyTrrpylsys 26
Db      128 CTCCTCGAGAGCTTTGCCGATCTCGACAGCTGGAGTCCCGCTGCGTGGTGGGCG 187
Qy      27 LeuAlaValAla1lelylsyrsProserGlyAspAspArgTyraSnglnPheHis1leArg 46
Db      188 CTGGCGAGCGACTTTCAGC-----AGCTGGCTGATGTTCTGCT 226
Qy      47 ArgPheGluAla1leuSnglnThrglylsyrsProThrsSerGlnleuSnglnPheAspTyr 66
Db      227 CATATTGAAAGTACTTAACCAAGGTAAAGTGAACAGAGAAATTAATGCTGCTGCTGG 286
Qy      67 GlyThrtThrsnCysThrsValGlyAspLeuValAspLeuLeu 80
Db      287 GCACAGAAAAACAAGCATCGGTGACCTTTTAAAGTCTTC 328

RESULT 13
US-10-198-846-1432
; Sequence 1432, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinhmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049

```

```

; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1432
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 2, 7, 217
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-1432

Alignment Scores:
Pred. No.: 0.0105 Length: 476
Score: 79.50 Matches: 24
Percent Similarity: 45.45% Conservative: 16
Best Local Similarity: 27.27% Mismatches: 37
Query Match: 15.38% Indels: 11
DB: 15 Gaps: 3

US-10-001-254-6 (1-98) x US-10-198-846-1432 (1-476)
Qy      2 TyrValArgCysLeuAsnValGlyLeu-----IleArglys 13
Db      154 TACTGAGGGCTTATCTCTGCGCTTGGCCATCCCTCCCACTTATTTCTGAAGAA 213
Qy      14 LeuSerAspPhe1leAspProGlnGlnGlyTrrpylsysLeuAlaValAla1lelylsys 33
Db      214 GTANCCCATTTTTCACATTCGAGAGGATGGGA-----GCCAGGCGCGCTCTGA 267
Qy      34 ProserGlyAspAspArgTyraSnglnPheHis1leArgArgPheGluAla1leuSngln 53
Db      268 ACTAGTGGATCCCGCGGCTGCAGAGAAATTCATCAAGCTTATGATACCGTCGACCTC 327
Qy      54 ThrglylsyrsProThrsSerGlnleuSnglnPheAspTyrGlyThrtThrsnCysThrsVal 73
Db      328 GAGGGG---GGGCGCGGTACCGACTTTGTTCCTTTAGAGAGGTTAATTGCGCGCTT 384
Qy      74 GlyAspLeuValAspLeuLeu 81
Db      385 GCGGTATCATGCTATAGCTGTT 408

RESULT 14
US-10-366-288-19
; Sequence 19, Application US/10366288
; Publication No. US20030216288A1
; GENERAL INFORMATION:
; APPLICANT: Powell, Douglas
; APPLICANT: Welch, Nadine S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: AIDS AND HIV-RELATED DISORDERS USING 1414, 1481, 1553,
; TITLE OF INVENTION: 34021, 1720, 1683, 1552, 1682, 1675, 12825, 9952, 5816,
; TITLE OF INVENTION: 10002, 1611, 1371, 14324, 126, 270, 312, 167, 326, 18926,
; TITLE OF INVENTION: 6747, 1793, 1784 OR 2045 MOLECULES
; FILE REFERENCE: MP102-0255PRNOMNIM
; CURRENT APPLICATION NUMBER: US/10/366,288
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 60/357,391
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/380,249
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: 60/391,306
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 60/406,297
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: 60/412,007
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: 60/417,508
; PRIOR FILING DATE: 2002-10-10

```


PRIOR APPLICATION NUMBER: 60/432,318
PRIOR FILING DATE: 2002-12-10
NUMBER OF SEQ ID NOS: 52
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19
LENGTH: 1782
TYPE: DNA
ORGANISM: Homo Sapien
US-10-366-288-19

Alignment Scores:
Pred. No.: 0.0844 Length: 1782
Score: 79.00 Matches: 22
Percent Similarity: 49.28% Conservative: 12
Best Local Similarity: 31.88% Mismatches: 19
Query Match: 15.28% Indels: 16
DB: Gaps: 2

US-10-001-254-6 (1-98) x US-10-366-288-19 (1-1782)

QY 45 IleaRgArpPhegluaLaLeuLeuGlnThr---GlyLySerProThrSerGluLeuLeu 63
Db 133 CTGCGGAAGATCAAGTCATGAGCGGCTGCAGCGGTGACATCAGCGGAGCTGCTG 192
QY 64 PheaApTrpGlyThrThrAsnCysThrValGlyAspleuValAspleuLeuIleGlnAsn 83
Db 193 TGGTGGTGGGCGCATGCGGCGGAGCCAGCCGTCAGCAACTTGTGACCTCTGTGCCGCTG 252
QY 84 GluphehealAProAlaSerLeuLeu----- 93
Db 253 GAGCTTACCGGGCTGCCAGATCATCTGAATGGAACCGGCTCTGAATCAGGTGT 312
QY 94 -----ProAspAlaVal 97
Db 313 CCCATTCCAGCCTTCCCTGACTGTG 339

RESULT 15
US-09-773-753-1
Sequence 1: Application US/09773753
Patent No. US20020099177A1

GENERAL INFORMATION:
APPLICANT: NI, JIAN

FENG, PING
MUZIO, MARTA
DIXIT, VISHVA M.

TITLE OF INVENTION: HUMAN IRAK-2
NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON

STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/773,753

FILING DATE: 02-Feb-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/980,060

FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:

NAME: STEFF, ERIC K.
REGISTRATION NUMBER: 36,688

REFERENCE/DOCKET NUMBER: 1488.1010000/EKS/AJK
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1806 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:
NAME/KEY: CDS

LOCATION: 34..1803
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-773-753-1

Alignment Scores:

Pred. No.: 0.086 Length: 1806
Score: 79.00 Matches: 22
Percent Similarity: 49.28% Conservative: 12
Best Local Similarity: 31.88% Mismatches: 19
Query Match: 15.28% Indels: 16
DB: Gaps: 2

US-10-001-254-6 (1-98) x US-09-773-753-1 (1-1806)

QY 45 IleaRgArpPhegluaLaLeuLeuGlnThr---GlyLySerProThrSerGluLeuLeu 63
Db 157 CTGCGGAAGATCAAGTCATGAGCGGCTGCAGCGGTGACATCAGCGGAGCTGCTG 216
QY 64 PheaApTrpGlyThrThrAsnCysThrValGlyAspleuValAspleuLeuIleGlnAsn 83
Db 217 TGGTGGTGGGCGCATGCGGCGGAGCCAGCCGTCAGCAACTTGTGACCTCTGTGCCGCTG 276
QY 84 GluphehealAProAlaSerLeuLeu----- 93
Db 277 GAGCTTACCGGGCTGCCAGATCATCTGAATGGAACCGGCTCTGAATCAGGTGT 336
QY 94 -----ProAspAlaVal 97
Db 337 CCCATTCCAGCCTTCCCTGACTGTG 363

Search completed: January 18, 2004, 09:05:28
Job time: 153.603 secs

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 18, 2004, 01:11:43 ; Search time 956.969 Seconds
(without alignments)
2488.940 Million cell updates/sec

Title: US-10-001-254-6
Perfect score: 517
Sequence: 1 TYVRCINVGIRKLSDPDP.....LLIQNEFFAPASLLLPDAPV 98

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 segs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O=/cgm2_1/USPTO.spool/US10001554/runat_16012004_152424_19723/asp_query.fasta_1.1109
-DB=EST -QPM=fastcap -SUFFIX=ret -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10001254.@CGN 1.1.4382.@runat_16012004_152424_19723 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em_estba:*
2: em_estchum:*
3: em_estcin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_juv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_rco:*
25: em_gss_pro:*
26: em_gss_pig:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	517	100.0	859	10	BG164491
2	509	98.5	811	10	BG164491
3	495	95.7	402	10	BE482619
4	485	93.8	719	10	BF669891
5	476	92.1	503	14	CA538859
6	476	92.1	598	13	BQ552228
7	476	92.1	610	10	BB660378
8	476	92.1	637	10	BB613447
9	476	92.1	638	14	BY721552
10	476	92.1	663	14	BY726858
11	476	92.1	676	10	BB613167
12	476	92.1	1161	11	AK0260397
13	476	92.1	2481	11	AK028837
14	476	92.1	2810	11	AK029028
15	472	91.3	575	4	BX522921
16	430.5	83.3	524	9	AW106160
17	415	80.3	541	10	BG691069
18	414.5	80.2	507	10	BB866698
19	393	76.0	453	10	BB860349
20	393	76.0	852	13	BU209111
21	389	75.2	600	9	AJ453616
22	389	75.2	670	9	AJ447581
23	369	71.4	265	9	AW436511
24	349	67.5	313	12	BM151935
25	343	66.3	629	9	AL647125
26	323.5	62.6	664	12	BQ35962
27	296	57.9	632	14	CA365604
28	289	55.9	733	14	CA474136
29	287	55.5	555	9	AW423082
30	270	52.2	520	10	BE132064
31	264	51.1	380	12	BI883790
32	258	49.9	505	10	BE482323
33	247	47.8	284	9	AA114228
34	247	47.8	540	9	AL699213
35	244	47.2	576	10	BF238344
36	237	45.8	781	10	BF723681
37	236	45.6	795	13	BQ293310
38	226	43.7	270	10	BE845841
39	204	39.5	480	9	AJ443799
40	193	37.3	888	10	BE687921
41	132	25.5	318	14	CB940191
42	124	24.0	595	28	AO601806
43	121.5	23.5	476	9	AV976730
44	121.5	23.5	509	9	AV995166
45	121.5	23.5	567	12	BP015417

ALIGNMENTS

RESULT 1
LOCUS BG164491
DEFINITION 602342026F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4452055 5', mRNA sequence.
ACCESSION BG164491
VERSION BG164491.1 GI:12671194
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 859)


```

DEFINITION 169463 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE482619
VERSION BE482619.1 GI:3602152
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 402)
Sonstegard, T., Capuco, A.V., White, J., Van Tassel, C.P., Connor, E.F.,
Cho, J., Sultana, R., Shade, L., Wray, J.E., Wells, K.D. and
Quackenbush, J.
TITLE Analysis of bovine mammary gland EST and functional annotation of
the Bos taurus gene index
JOURNAL Mamm. Genome 13 (7), 373-379 (2002)
MEDLINE 22135956
PUBMED 12140684
COMMENT Contact: Sonstegard TS
USDA, ARS, Beltsville Agricultural Research Center
Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
Tel: 301 504 8416
Fax: 301 504 8414
Email: tads@ps1.barc.usda.gov
Single pass sequencing. Bases called and alt. trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTGCTACGACG
Plate: 11 row: F column: 7
Seq primer: ATTAGTGACATGATAG.
FEATURES
source
1..402
location/Qualifiers
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="BARC 5BOV"
/note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."
BASE COUNT 114 a 95 c 88 g 105 t
ORIGIN
Alignment Scores:
Pred. No.: 5.03e-59 Length: 402
Score: 495.00 Matches: 92
Percent Similarity: 97.96% Conservative: 4
Best Local Similarity: 93.88% Mismatches: 2
Query Match: 95.74% Indels: 0
Gaps: 0
DB: 10
US-10-001-254-6 (1-98) x BE482619 (1-402)
QY 1 ThTyrValArgCysLeuAnValGlyLeuIleArgLysLeuSerAspPheIleAspPro 20
Db 88 ACATACGTCGCGCTCCATGCTTGAATTGAAAGCTGTCGATTTTATGATCTT 147
QY 21 GlnGluGlyTTrpLysLysLeuAlaValAlaIleLysLysProSerGlyAspAspArgTyr 40
Db 148 CAGAAGAAGATGAGAAAGTTAGCGATGCGCTATTAAAAACCATCTGATGATGATAC 207
QY 41 AasnGlnPheHisIleArgArgPheGlnAlaLeuLeuGlnThrGlyLysSerProThrSer 60
Db 208 AATCAGTTTCACATAGAAGATTGGAAGCATTAATGCAAAATTGAAAAAGCCACGCTG 267
QY 61 GlnLeuLeuPheAspTTrpGlyThrThrAsnGlyThrValGlyAspLeuValAspLeuLeu 80
Db 268 GAGTTACTGTTGTGCTGGGGGACCAAAATTGACAGTTGGTGTATCTTGGATATTTTG 327

```

```

QY 81 IlleGlnAsnGlnPheAspPheAlaProAlaSerLysLeuLeuProAspAlaValPro 98
Db 328 GTCCAAATATAGTTTTTTTGGCCCTGCAAGCTTTTGTCTACAGATGCTGTAACC 381
RESULT 4
LOCUS BF696981
DEFINITION BF696981 719 bp mRNA linear EST 22-DEC-2000
602130160F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4287014 5',
mRNA sequence.
ACCESSION BF696981 GI:11982389
VERSION BF696981.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 719)
NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gsabbs-rc@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM124 row: 9 column: 15
High quality sequence stop: 632.
FEATURES
source
1..719
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4287014"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NIH_MGC_56"
/note="Organ: Brain; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggcgccctcgccg); Site 2: SfiI (ggccatcagccg);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCATTATGCGC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGAGCGCGCGCATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 216 a 145 c 163 g 195 t
ORIGIN
Alignment Scores:
Pred. No.: 2.87e-57 Length: 719
Score: 485.00 Matches: 97
Percent Similarity: 97.00% Conservative: 0
Best Local Similarity: 97.00% Mismatches: 1
Query Match: 93.81% Indels: 2
Gaps: 0
DB: 10
US-10-001-254-6 (1-98) x BF696981 (1-719)
QY 1 ThTyrValArgCysLeuAnValGlyLeuIleArgLysLeuSerAspPheIleAspPro 20
Db 92 ACATATGCGCGCTCCATGCTTGAATTGAAAGCTGTCGATTTTATGATCTT 151
QY 21 GlnGluGlyTTrpLysLysLeuAlaValAlaIleLysLysProSerGlyAspAspArgTyr 40
Db 152 CAGAAGAAGATGAGAAAGTTAGCGATGCGCTATTAAAAACCATCTGATGATGATAC 211
QY 41 AasnGlnPheHisIleArgArgPheGlnAlaLeuLeuGlnThrGlyLysSerProThrSer 60

```


Oy		41	Aasnlnphhsllleayagapghglualaleuenglntbrgilyvsasrprothser	60
Db		332	AATCAGTTCATATAAAGAGATTGAAGCCCTTACTTGACACGGGAAAGACCACCTGT	391
Oy		61	GltucleutPheasPTPGlyThrThrsAntCythrValGlYAspleauValAspleuLeu	80
Db		392	GAATCGCTGTTGACTGGGGACCAACGACACTGCACAGTGGCGACTTGTGAGACTTCTG	451
Oy		81	IleglinsngluipheplialproalaserleuleuProaaplalavalpro	98
Db		452	GTCGACGATTGAGCTGTGTGCCCGCCGCACTCTCCTGCTCGCGAGATGCCGTTCCC	505
RESULT 8				
Bb613447				
LOCUS				
DEFINITION	Bb613447	637 bp	mRNA	linear EST 26-Oct-2001
ACCESSION	Bb613447			
VERSION	Bb613447			
KEYWORDS	Bb613447.1	GI:16454055		
SOURCE				
ORGANISM	Mus musculus (house mouse)			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciuromorphi; Muridae; Murinae; Mus.			
TITLE	musculus cDNA clone 4732482P03 5', mRNA sequence.			
JOURNAL	Bb613447			
COMMENT	Bb613447.1 GI:16454055			
	EST.			
	Mus musculus			
	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciuromorphi; Muridae; Murinae; Mus.			
	1 (bases 1 to 637)			
	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.			
	Hikemoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Komno,H., Kouda			
	Koyama,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,			
	Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki			
	D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,			
	Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,			
	Muramatsu,M., and Hayashizaki,Y.			
	Riken Mouse ESTs (Arakawa,T., et al. 2001)			
	Unpublished			
	Contact: Yoshihide Hayashizaki			
	Laboratory for Genome Exploration Research Group, RIKEN Genomic			
	Sciences Center(GSC), Yokohama Institute			
	The Institute of Physical and Chemical Research (RIKEN)			
	1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan			
	Tel: 81-45-503-9222			
	Fax: 81-45-503-9216			
	Email: genome-res@gsc.riken.go.jp,			
	URL:http://genome.gsc.riken.go.jp/			
	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh			
	M., Komno,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.			
	Normalization and subtraction of cap-trapper-selected cDNAs to			
	prepare full-length cDNA libraries for rapid discovery of new			
	genes. Genome Res. 10 (10), 1617-1630 (2000)			
	wegli,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,			
	Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunaga			
	S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A., and			
	Hayashizaki,Y.			
	RIKEN integrated sequence analysis (RISA) system--384-format			
	sequencing pipeline with 384 multicapillary sequencer. Genome Res.			
	10 (11), 1757-1771 (2000)			
	Komno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara			
	,Y., and Hayashizaki,Y.			
	Computer-based methods for the mouse full-length cDNA			
	encyclopedia: real-time sequence clustering for construction of a			
	nondundant cDNA library. Genome Res. 11 (2), 281-289 (2001)			
	Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamana,I., Aizawa			
	K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., and			
	Hayashizaki,Y.			
	Computational Analysis of Full-Length Mouse cDNAs Compared with			
	Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)			
	Please visit our web site (http://genome.gsc.riken.go.jp) for			
	further details.			
	e mouse tissues.			
FEATURES				
SOURCE				
	Location/Qualifiers			
	1..637			
	/organism="Mus musculus"			
	/mol_type="mRNA"			
	/strain="C57BL/6J"			

```

/db xref="taxon:10090"
/cloae="A.732482P03"
/sex="mixed"
/tissue_type="skin"
/dev stage="10 days neonate"
/lab host="DH10B"
/clone_1lb="RIKEN full-length enriched, 10 day neonate skin"
/note="Site_1: Salt; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAAAGCTCTTTTTCCTTTTTTTTNN 3'], cDNA was prepared by using triphase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGCGATTATTAATAATCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I"
```

BASE COUNT 162 a 185 c 173 g 117 t

ORIGIN

Alignment Scores:

Pred. No.:	4,41e-56	Length:	637
Score:	476.00	Matches:	89
Percent Similarity:	95.92%	Conservative:	5
Best Local Similarity:	90.82%	Mismatches:	4
Query Match:	92.07%	Indels:	0
DB:	10	Gaps:	0

US-10-001-254-6 (1-98) x BB613447 (1-637)

Oy	1	ThrTyValAlrgCySleAsnVnAlglyLeuIlleArglySleuSerAppheIlleasPro	20
Db	193	ACATACATACGCAACCCTTAATGTGGGAGACTTATGAGAAGCTGTGCGATTTTATGATGCCT	252
Oy	21	GlnGlubLyrTrpIysLySleuAlaValAlaIlleYelYspProSerGIYaaspARgTYr	40
Db	253	CAGAAGGGTGGAAGAAATTAGACAGTAGCTATCAAAGCCGTCCGCCAGACAGATAC	312
Oy	41	AsnGlnPheHisIleArGrpPheGlnAlaleuLeuGlnThrGIylYSerProThSer	60
Db	313	AATAGATTCATNAAAGAGATTGGAAGCCTTACTTCAGACCCGGAAGGCCAACCTGT	372
Oy	61	GluLeuLeuPheApPTpgIlyThrThraScyThrValGIylaPleuValaSPleuLeu	80
Db	373	GAACTGCTGTTGACTGGGCGACACCAACAACGTGCACAGTTGGCGACCTGTGGATCTACTG	432
Oy	81	IleGlnsnnglunPhePheAlaProAlaSerLeuLeuLeuProaPaAlaValPro	98
Db	433	GTCACAGATTAGCTGTTGCCGCCGCCCATCTCTGCTGCCGGANGCCGTTCCC	486

RESULT 9
BY721552 LOCUS BY721552 RIKEN full-length enriched, adult male diencephalon Mus musculus cDNA clone 9310209D03 5', mRNA sequence.

ACCESSION BY721552
VERSION BY721552.1 GI:27134669
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 638)
Okazaki,Y., Furuno,M., Kaukawa,T., Adachi,J., Bono,H., Kondo,S., Nishida,I., Osato,R., Saito,R., Suzuki,H., Yamataka,I., Kiyosawa,H.

Bettsel, K.W., Blake, J.A., Brad, D., Brusic, V., Choitha, C., Cochran,
 L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest,
 A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
 Gough, J., Grimonod, S., Gustinchich, S., Hirokawa, N., Jackson, I.U.,
 Jarvis, E.D., Karai, A., Kawaji, H., Kawasawa, Y., Kedierski, R.M.,
 King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lemhar, B., Lyons,
 P.A., Maglocz, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pettes, G.,
 Peccole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D.,
 Ramachandran, S., Ravasi, T., Reed, U.C., Reed, D.U., Reid, J., Ring,
 B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou,
 M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale,
 R.D., Tomlita, M., Verrardo, R., Wagner, L., Walshester, C., Wang, Y.,
 Watanabe, Y., Wells, C., Wilming, L.G., Wynchaw-Boris, A., Yanagisawa,
 M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
 Caminici, P., Hayatsu, N., Hirozane-Kishikawa, T., Komno, H., Nakamura,
 M., Sakakuma, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aikawa, K.,
 Arikawa, T., Fukuda, S., Hara, A., Hashizume, W., Imcanti, C., Ishii,
 Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Saeki, D., Shibara,
 K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
 E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

JOURNAL
 MEDLINE
 PUBMED
 12354683
 12466851

TITLE
 COMMENT
 Contact: Yoshinhide Hayashizaki

Kono, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasahara, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN Integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE
PUBMED
20530913
11076861

REFERENCE
AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aikawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kanakawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Knehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hochmann, M., Hume, D. A., Kamliya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyokawa, K., Wang, K. H., Weitz, C., Whitaker, C., Wilmink, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohetsuki, S., and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

JOURNAL MEDLINE
PUBMED
21085660
11217851

REFERENCE
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL MEDLINE
PUBMED
11217851

REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hatanaka, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirose, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasuwa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraishi, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

JOURNAL MEDLINE
PUBMED
20530913
11076861

REFERENCE
AUTHORS

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1. 2481
/organism="Mus musculus"

FEATURES
Source

/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:4732460109"
/db_xref="taxon:10090"
/clone="4732460109"
/tissue_type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/cds_start="10 days neonate"
116_1495
/note="unlabeled protein product; interleukin-1 receptor-associated kinase 4 [Mus musculus]
putative"
/codon_start=1
/protein_id="BAC26146.1"
/db_xref="GI:26324784"
/translation="MANKPLTPSYIRNANGLIRKSDPFDPOGKRLAVALIKPSG
DDVYNDPHIRFALLQTKSPFCELLFDWGTTCVGDVLDVQLFAPATLLP
DAVQTVKSLPPEEAATVQTHPCQCKRTSYMPKLEHSCPEPSSPPDKRSVS
SDPRFHSFHEKSLITNNDPEQASGRMGEGGVYKCVNNTIYAVKGLGM
VELSTELKQPOEIKVATCOHEMLVELLGGSSDNLCLVYAVPNSLDRISC
LDGTPPLSMTRCKVAGTANGIRFLHNNHHRDLSANILDKPTAISDGLAR
ASRLAQTVTSRIVGTIYMAPEALRGELTPSDIYSFVVLLELTGLAAVDENE
PQLLDIKERIEDEKTIEDYTEKMSDADPASVEAMYSAASQCLHEKORRDIAY
QQLLOEKSA"

CDS

BASE COUNT 650 a 592 c 634 g 605 t

ORIGIN

Alignment Scores:
Pred. No.: 2,986-55 Length: 2481
Score: 476.00 Matches: 89
Percent Similarity: 95.92% Conservative: 5
Best Local Similarity: 90.82% Mismatches: 5
Query Match: 92.07% Indels: 0
DB: 11 Gaps: 0

US-10-001-254-6 (1-98) x AK028837 (1-2481)

QY 1 ThrTyrValArgCysLeuAsnValGlyLeuLeuArgGlyLeuSerAspPheIleAspPro 20
DB 140 ACATACATACGCAACCTTAATGAGGATCTTATGGAAGCTGTGGATTTATGATCCT 199
QY 21 GlnGlnGlyTTrpLysLysLeuAlaValAlaIleLysLysProSerGlyAspAspArgTyr 40
DB 200 CAGAAGGGTGAAGAATTAAGACATGATCAATAAAGCCGTCGGCAGCAGACATAC 259
QY 41 AsnGlnPheHisIleArgArgPheGluAlaLeuLeuGlnThrGlyLysSerProThrSer 60
DB 260 AATCAGTTCATTAAGAGATTCGAAGCCTTACTTACAGCCGGGAAGACCCACCTGT 319
QY 61 GlnLeuLeuPheAspTrpGlyThrThrAsnCysThrValGlyAspLeuValAspLeuLeu 80
DB 320 GAACCTGCTGTATACGCGGACACGACGATGCAAGTGGCGACCTGTGGATCTACCTG 379
QY 81 IleGlnAsnGlnPhePheAlaProAlaSerLeuLeuLeuProAspAlaValPro 98
DB 380 GTCCAAATTCAGCTGTTTGGCCCCCGCACCTCTCTGCGGAGATGCCGTTCCC 433

RESULT 14
AK029028 2810 bp mRNA linear HTC 05-DEC-2002
LOCUS Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched
DEFINITION library, clone:4732462P03 product:interleukin-1 receptor-associated
kinase 4 [Mus musculus], full insert sequence.

ACCESSION AK029028
VERSION AK029028.1 GI:26081132
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batilov, S., Casavant, T., Pleschmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikiido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staudli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Botelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Guestrinich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schombach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weltz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S., and Hayashizaki, Y.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085660
 PUBMED 11217851
 REFERENCE 5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 MEDLINE 12028101
 PUBMED 12028101
 REFERENCE 6
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyata, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp)

COMMENT URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216
 CNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse libraries.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 URL: http://location/Qualifiers

FEATURES
 source
 1. 2810
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:4732482P03"
 /db_xref="taxon:10090"
 /clone="4732482P03"
 /tissue_type="skin"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="10 days neonate"
 176..1556
 /note="Interleukin-1 receptor-associated kinase 4 [Mus musculus]"
 putative

misc_feature
 176..1556
 /note="Interleukin-1 receptor-associated kinase 4 [Mus musculus]"
 putative

BASE COUNT 738 a 677 c 715 g 680 t

ORIGIN

Alignment Scores:
 Pred. No.: 3,546-55 Length: 2810
 Score: 476.00 Matches: 89
 Percent Similarity: 95.92% Conservative: 5
 Best Local Similarity: 90.82% Mismatches: 4
 Query Match: 92.07% Indels: 0
 DB: 11 Gaps: 0

US-10-001-254-6 (1-98) x AK029028 (1-2810)

QY 1 ThrTtVtAlArCyLeuAbnVaIglYleuIleArGtYleuSerAspHeIleAspPro 20
 DB 200 ACATTCATRCGACACCTTATATGCGGATCTTACGAACTCTCGGATTTATTCATCT 259

QY 21 GtngtuglYtRlyeLysLeuAlaValAlaIleLysLysProSerGtYAspAspArgTy 40
 DB 260 CAAGAAGGTGAGAAATTAAGCAGTATCAAAAGCCGTCCGGCAGACAGATAC 319

QY 41 AaNgInPheHsIleArGrhPheGtLAlaLeuLeuGlnThGtYlYsSerProThrSer 60
 DB 320 AATCGATTCATATTAAGATTTCAAGCCTTACAGACCGGAGAGGCCCACTCT 379

QY 61 GluLeuLeuPheAspTtRgLYtThrThrAsnCySthrVaIglYAspLeuValAspLeu 80
 DB 380 GAACGTGCTTTGACTGGGCGACCAAGACCTTCAAGCTTGGGAGCTTGTGATCTAC 439

QY 81 lIeGInAsngLpHePheHsIleArGrhPheGtLAlaLeuLeuGlnThGtYlYsSerPro 98
 DB 440 GTCAGATTAAGCTGTTTCCCGGCACTCTCCGCTGTCGAGTCCGATCC 493

RESULT 15
 BX522921 standard; RNA; EST; 575 BP.
 ID BX522921
 XX BX522921;
 AC BX522921;
 XX BX522921.1
 SV BX522921.1
 XX
 DT 27-MAY-2003 (rel. 75, Created)
 DT 27-MAY-2003 (rel. 75, Last updated, Version 1)
 XX
 DE RZPD Mus musculus cDNA clone IMAGE9981225504 = IMAGE:2225445 5' EST.
 XX EST; expressed sequence tag.
 XX

OS Mus musculus (house mouse)
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
 CC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 XX
 RN [1]
 RP 1-575
 RA Heil O., Ebert L., Neubert P., Peters M., Radelof U., Schneider D.,
 RA Korn B.;
 RT
 RL Submitted (28-MAY-2003) to the EMBL/GenBank/DBJ databases.
 RL RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH im Neuenheimer
 RL Feld 580, D-69120 Heidelberg, Germany
 XX
 CC RZPD; IMAGE9981225504.
 CC RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
 CC Mouse Unigeneset - RZPD2 (RZPDLIB No.981)
 CC http://www.rzpd.de/Clonedata/cgi-bin/showlib.pl.cgi?response?libNo=981
 CC Contact: Ina Rolfs
 CC RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 CC Heubnerweg 6, D-14059 Berlin, Germany
 CC Tel: +49 30 32639 101
 CC Fax: +49 30 32639 111
 CC www.rzpd.de
 CC This clone is available royalty-free from RZPD;
 CC contact RZPD (clone@rzpd.de) for further information.
 CC Seq primer: sugf, Primer sequence: CTTGCTCTTAAAGCTGCG
 XX
 FH Key
 FH Location/Qualifiers
 FT
 FT source
 FT 1. 575
 FT /db_xref="taxon:10090"
 FT /note="1st strand cDNA was primed with an oligo(dT) primer
 FT [ATGGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
 FT ligated to a DraIII adaptor [TGTGGCTTACTG], digested and
 FT cloned into distinct DraIII sites of the pME18S-FL3 vector
 FT (5' site CACTGTGG, 3' site CACCATGTG). XhoI should be used
 FT to isolate the cDNA insert. Size selection was performed to
 FT exclude fragments <1.5kb. Library constructed by Dr. Sumio
 FT Sugano (University of Tokyo Institute of Medical Science).
 FT Custom primers for sequencing: 5' end primer
 FT CTTGCTCTTAAAGCTGCG and 3' end primer
 FT CGACCTGACCTCGACGACA. REFERENCES: Suzuki, Y., Yoshitomo,
 FT K., Maruyama, K., Suyama, A., and Sugano, S. Construction
 FT and characterization of a full length-enriched and a 5' end
 FT enriched cDNA library. Gene 200, 149-156, 1997. Sasaki, Z.,
 FT Suzuki, Y., Matanabe, M., Imai, J., Shibui, A., Yoshida,
 FT K., Hata, H., Yamaguchi, R., Tateyama, S., and Sugano, S.
 FT Construction of mouse full length-enriched cDNA libraries
 FT by oligo-capping. DNA Research, submitted."
 FT /organism="Mus musculus"
 FT /clone="IMAGE9981225504"
 FT /clone_lib="Sugano mouse embryo mewa"
 FT /dev_stage="embryo, 14 dpc"
 FT /lab_host="DH10B"
 XX
 SQ Sequence 575 BP; 132 A; 172 C; 155 G; 114 T; 2 other;
 Alignment Scores:
 Pred. No.: 1.39e-55 Length: 575
 Score: 472.00 Matches: 88
 Percent Similarity: 95.92% Conservative: 6
 Best Local Similarity: 89.80% Mismatches: 4
 Query Match: 91.30% Indels: 0
 DB: 4 Gaps: 0
 US-10-001-254-6 (1-98) x BK522921 (1-575)
 QY 1 ThrtYrValArgCysLeuAenValGlyLeuIleArdIysLeuSerAaspHeIleAaspPro 20
 DB 240 ACATACATACGCAACCTTAATGGGATCCTTAAGAAAGCTGCGAATTATATGATCCT 299
 QY 21 GlnGlnGlyTTrpLysLysLeuAlaValAlaIleLysLysProSerGlyAaspAparGlyTyr 40

DB 300 CAAAGAGGTTGGAAGAAATTAGCAGTACTATCAAAAAAGCCGTCGGCGACGACGATAC 359
 QY 41 AaGlnbHeHsIleAargAphEgLuAlaLeuGlnTrhGlyLysSerProThSer 60
 DB 360 AATCAGTTCCATTAAGAGAGATTGGAAGCCCTTACTTCAAGCCGGGAAGCCCACTGT 419
 QY 61 GluLeuLeuPheAePTPGlyThrThrAsnCysThrValGlyAaPLeuValAaPLeuLeu 80
 DB 420 GAACTGCTGTTTGACTGGGCGACACGAACTGCACAGTTGGCGACCTTGATCTACTG 479
 QY 81 IlEglnAaGlnPhePheAlaProAlaSerLeuLeuLeuProAaPAlaValPro 98
 DB 480 GTCCAGATTGAGCTGTTGCCCGCCGACCTCTGCTGCCGGAAGCCGTTCCC 533

Search completed: January 18, 2004, 06:06:40
 Job time : 960.969 secs

BEST AVAILABLE COPY

XX Key Location/Qualifiers
FH 1..1383
FT CDS /tag=a
FT /product= "Human IRAK4"
PN MO200240680-A2.
XX
XX 23-MAY-2002.
XX
XX 15-NOV-2001; 2001WO-US44844.
XX
XX 17-NOV-2000; 2000HS-0715893.
PR 29-JUN-2001; 2001US-301889P.
XX
PA (BURN-) BURNHAM INST.
PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W,
PI Stenner-Jewen F;
DR WPI; 2002-500222/53.
XX P-PsDB; AA024859.
XX
PT New polypeptide comprising a death domain or death effector domain,
PT useful for discovery of drugs that suppress infection, inflammation,
PT allergy, sepsis, autoimmunity, allograft rejection and other diseases
PT -
PS Claim 19; Page 180-182; 209pp; English.

The invention relates to an isolated polypeptide comprising a death domain (DD), death effector domain (DED) or NB-ARC domain. The invention is useful for identifying a binding agent, preferably a protein or a drug that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC domain from DAP3, IRAK4, CTSD (Chlamydia trachomatis DP protein), DED4 or NIDD (NGFR-interacting Death Domain), with a candidate binding agent and detecting the association of the domain and the candidate binding agent, by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or chemical crosslinking, nuclear magnetic resonance (NMR), mass spectroscopy (MS) and FPLC. The invention is useful for modulating the level of a cell process such as cell proliferation, cell adhesion, cell stress responses, responses to microbial infection and B cell immunoglobulin class switching, in particular apoptosis within a cell. Antibody specifically reactive with CTSD DD of C. trachomatis, C. muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the CTSD DP protein is useful for detecting a Chlamydia infection. The invention is useful for modulating the activity of oncogenic proteins, for treating a pathology caused by the oncogenic proteins and for treating bacterial infections by modulating the activity of bacterial proteins. The protein and antibody specific for it are useful for discovery of drugs that suppress infection, inflammation, allergy, sepsis, autoimmunity, allograft rejection and other diseases. The protein is useful for treating immune-based pathologies, pathologies associated with cell division, inflammatory diseases such as sepsis, fibrosis, arthritis, graft versus host disease. The invention is used in antisense therapy and gene therapy. The present sequence is human IRAK4 gene.

SQ Sequence 1383 BP; 463 A; 243 C; 283 G; 394 T; 0 other;

Alignment Scores:

Pred. No.:	Length:
6,59e-234	1383
Score: 2386.00	Matches: 460
Percent Similarity: 100.00%	Conservative: 0
Best Local Similarity: 100.00%	Mismatches: 0
Query Match: 100.00%	Indels: 0
DB: 24	Gaps: 0

US-10-001-254-16 (1-460) x AAD40079 (1-1383)

CY 1 MetAanlySProllePhProserThrrYrValAlarGyleuanvalGlylenlllearg 20
Db 1 ATGAACCAACCCATATCACCATATCATATGTGGCTCGCTCATATGTTCAGTAAATTAGG 60

QY	2	LysLeuSerAspPheIleAspProGlnGlyTyrPheLysLeuAlaValIleLys	40
Db	61	AAAGCTGCAGATTTTATTGATCCCGAAGAGATGAGAGAGTTAGCTGTCGATTAAA	120
QY	41	LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgArgPheGlnAlaLeu	60
Db	121	AAACCATCTGGTGTAGTATGATACATCACTGTTTCACTAAGAGATTTTGAAGCATTTACTT	180
QY	61	GlnThrGlyLysSerProThrSerGlnLeuLeuPheAspTrpGlyThrThrAsnCysThr	80
Db	181	CAAACTGGAAAAGATCCCACTTCGAAATTACGTGTTACTCGGGGACACAAATTTGACA	240
QY	81	ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPheAlaProAlaSerLeuLeu	100
Db	241	GTGGTGATCTTGGAGATCTTTTATGTCAAATGAAATTTTGTCTCTCGGAGCTTTTGG	300
QY	101	LeuProAspAlaValProLysThrAlaAsnThrLeuProSerLysGlnAlaIleThrVal	120
Db	301	CTCCAGATGGTGTCTCCAAAACGCTAATACATCACTCTTAAGAAAGCATAAACGTT	360
QY	121	GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn	140
Db	361	CAGCAAAAACAGATGCCCTTCTGTGACAAAGACAGGACATTGAGAACACCTGTGCAGAA	420
QY	141	LeuGlnGlnSerTyrMetProProAspSerSerSerProGlnAsnLysSerLeuGlnVal	160
Db	421	CTTAAGACAAACATAATATGACCACCTGACTCTCCAAATCCAGAAAATTAAGTTTGAAGTT	480
QY	161	SerAspThrArgPheHisSerPheSerPheTyrGlnLeuLysAsnValThrAsnAsnPhe	180
Db	481	AGATATACAGCTTTTCAACAGTTTTCATTTTATGAAATGAAGATGTACAAATATCACTT	540
QY	181	AspGluArgProIleSerValGlyGlyAsnLysMetGlyGlyGlyGlyPheGlyValVal	200
Db	541	GATACAGACACCATTTCTGTTGGTGTAATAAATGGAGAGGAGGATTTGGAGTTGTA	600
QY	201	TyrTyrGlyTyrValAsnAsnThrThrValAlaValLysLysLeuAlaIleMetValAsp	220
Db	601	TATAAAGGCTACGTAATTAACACAACTGTGGCAGTGAAGAGCTTGACGACATGTGGAC	660
QY	221	IleThrThrGlnGlnLeuLysGlnGlnPheAspGlnGlnIleLysValMetAlaLysCys	240
Db	661	ATTACTACTGAGAACTGGAACACGACATTTGATCAAGAAATTAAGTAATGGCAAAAGTT	720
QY	241	GlnHisGlnAsnLeuValGlnLeuLeuGlyPheSerSerAspGlyAspAspLeuCysLeu	260
Db	721	CAACATGAAAACTAGTAGAACTCTGTTCTCAAGTACATGAGATGACCTCTGCTTA	780
QY	261	ValTyrValTyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThr	280
Db	781	GTAATATGTTTACATGCTCAATATGTTCACTGTTCAACAGACTCTTCTTGGAGTGACT	840
QY	281	ProProLeuSerTyrHisMetArgCysLysIleIleIleGlnGlyAlaIleAsnGlyIleAsn	300
Db	841	CCACCACCTTCTTGCCACATGAGATGCAAGATTGCTCAGGGTGCAGCTAATGGCATCAAT	900
QY	301	PheLeuHisGlnAsnHisIleHisIleArgAspIleLysSerAlaAsnIleLeuLeuAsp	320
Db	901	TTTCTACATGAAATATCATCATATTCATAGAGATTTAAAGTGCAAAATATCTTACTGGAT	960
QY	321	GlnAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGlnLysPheAla	340
Db	961	GAACCTTTTACGTGTAATAATATCGACTTTGGCCTTGCAAGGCGCTTGAGAAATTTGGC	1020
QY	341	GlnThrValMetThrSerArgIleValGlyThrThrAlaTyrMetAlaProGlnAlaLeu	360
Db	1021	CAGACAGTCATGACTAGAGATGTGGAAACAACAGCTTAATATGGACACAGAAAGCTTTG	1080
QY	361	ArgGlyGlnIleThrProLysSerAspIleTyrSerPheGlyValValLeuLeuGlnIle	380
Db	1081	CGTGAGAAATTAACCCCAATCTGATATTTACAGCTTTGGTGTGTGTTTACTAGAAATA	1140
QY	381	IleThrGlyLeuProAlaValAspGlnHisArgGlnProGlnLeuLeuAspIleLys	400

Dd		1141	ATTAATGGA	CTTCCAGCTGTGGATGAACACCCTGAACCTCAATTTCGTACATTTTAA	12006
Oy		401	GUGUUII	EGUAAPCUGUULysThrIIeguaSPYrIIeaSPYLysWeeAsnaSP	420
Dd		1201	GAA	GAATAATTGAAGTGAAGAAAAGACAATTGCMAAGATVATATGATTAATAAAGATGAATCAT	1266
Oy		421	AlaaSPsrTrHrSeValGUlaMeTySerValAlaaSerGlnCyaleuHiSgluLys	440	
Dd		1261	GCTGATTC	CACCTCGATTGAGTGAAAGCTATGACTCTGTGTGTAGTCATCAATGTCGTGATGAAAAA	1320
Oy		441	LysaSenLysArGPProASPll	eLysLysValGInGlnLeuLeuGInGluMetThraLaSer	460
Dd		1321	AAAAATA	GAGACGCCAGACATTTAGAGAAGTTCAACAGCTGCTGCAGAGATGACAGCTTCT	1380
RESULT 2					
AAD10197	ID	AAD10197	standard;	cDNA; 1383 BP.	
XX	AC	AAD10197;			
XX	DT	24-SEP-2001	(first entry)		
DE	XX	Human interleukin-1 (IL-1) receptor-associated kinase (IRAK)-4 cDNA.			
KM	KM	Human, interleukin-1 receptor associated kinase-4; IRAK-4; cytosolic;			
KM	KM	IL; antibacterial; antiinflammatory; ophthalmological; vasotropic; OP;			
KM	KM	chronic obstructive pulmonary disease; neuroprotective; chronic cough;			
KM	KM	adult respiratory disease syndrome; pulmonary fibrosis; asthma; ARDS;			
KM	KM	interstitial lung disease; allergic rhinitis; transplant rejection;			
KM	KM	autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;			
KM	KM	multiple sclerosis; diabetes; cancer; solid tumour; lymphoma; stroke;			
KM	KM	cardiovascular disease; atherosclerosis; neurodegenerative disease;			
KM	KM	sepsis; osteoarthritis; rash; osteoporosis; psoriasis; dermatitis;			
KM	KM	inflammatory bowel disease; gout; Crohn's disease; ulcerative colitis;			
KM	KM	Behcet's syndrome; ankylosing spondylitis; IL-1 receptor/Toll receptor;			
KM	KM	sarcoidosis; transgenic animal; SE.			
OS	XX	Homo sapiens.			
FT	FT	Key	Location/Qualifiers		
FT	FT	CDS	1..1383		
FT	FT		/tag= a		
FT	FT		/product= "Human IRAK-4"		
XX	PN	MO200151641-A1.			
XX	PD	19-JUL-2001.			
XX	PF	12-JAN-2001; 2001WO-US01171.			
XX	PR	13-JAN-2000; 2000US-0176395.			
XX	PA	(TULIA-) TULARIX INC.			
P1	P1	Weesche H, Li S;			
DR	XX	WPI; 2001-451860/48.			
DR	XX	P-PADB; AAE05398.			
XX	PT	Novel human interleukin-1 receptor associated kinase polypeptide,			
XX	PT	useful for identifying modulators of the polypeptide for treating gout,			
XX	PS	asthma, allergic rhinitis, multiple sclerosis and skin cancer -			
XX	PS	Claim 7; Fig 2; 89pp: English.			
CC	CC	The present sequence is a cDNA encoding human interleukin (IL)-1 receptor			
CC	CC	associated kinase (IRAK)-4. IRAK associate with activated IL-1, IL-18			
CC	CC	and other receptors and act to transduce signals originating from the			
CC	CC	activated receptors, ultimately leading to a variety of downstream			
CC	CC	effects such as nuclear factor (NF)-kappaB activation. The IRAK-4			
CC	CC	inhibitors are useful for treating inflammatory diseases such as			
CC	CC	pulmonary diseases and diseases of the airway (e.g., adult respiratory			

	disease syndrome (ARDS), chronic obstructive pulmonary disease (COPD),
CC	pulmonary fibrosis, interstitial lung disease, asbestosis, chronic cough,
CC	or allergic rhinitis), transplant rejection, autoimmune diseases (e.g.,
CC	rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis
CC	or diabetes), cancer (e.g., solid tumours, skin cancer or lymphoma),
CC	cardiovascular diseases (e.g., stroke and arteriosclerosis), diseases
CC	of the central nervous system (e.g., neurodegenerative disease), CD4
CC	mediated sepsis, non-CD4 mediated sepsis, osteoarthritis, osteoporosis,
CC	psoriasis, diseases of the skin (e.g., rash, contact dermatitis, atopic
CC	dermatitis), inflammatory bowel disease (e.g., Crohn's disease and
CC	ulcerative colitis), Behcet's syndrome, ankylosing spondylitis, gout,
CC	sarcoidosis and ophthalmic diseases and conditions. The inhibitors of
CC	IRAK-4 activity or expression are used to inhibit signal transduction
CC	resulting from the activation of an interleukin-1 receptor (IL-1R)/Toll
CC	receptor in a cell. They also inhibit the activation of a transcription
CC	factor that activates NF-kappaB in the cell. IRAK-4 is used to create a
CC	nonhuman transgenic animal which is useful for testing the function of
CC	IRAK-4 in vivo, to generate models for the study of inflammatory
CC	disorders and conditions and for the development of potential treatments
CC	for IRAK-4-related inflammatory diseases and conditions. IRAK-4 sequences
CC	are also used in gene therapy and in antisense therapy.
XX	
SQ	Sequence 1383 BP; 463 A; 244 C; 283 G; 393 T; 0 other;
	Alignment Scores:
Pred. No.:	1,69e-233 Length: 1383
Score:	2382.00 Matches: 459
Percent Similarity:	99.78% Conservative: 0
Best Local Similarity:	99.78% Mismatches: 1
Query Match:	99.83% Indels: 0
Dbl:	22 Gaps: 0
US-10-001-254-16 (1-460) x AMD10197 (1-1383)	
OY	1 MetAsnLysProIleThrPseSerThrLyArgCyLeuAnaValGlyLeuIleArg 20
Dbl	1 ATAAACAACCAATTAACACCATCAATAATAGTGCGCTGCCAATGTGGACTTAATAGG 60
OY	21 LysLeuSerAspPheIleAspProGlnGluGlyTrpValSylLeuAlaValAlaIleLys 40
Dbl	61 AACCTGCAGATTATTTATGATGCTTCAGAAGATGAAGAAGTAGCTAGCATTAATAA 120
OY	41 LysProSerGlyAspAspArgTyraGlnPheHisIleArgArgPheGluAlaLeuLeu 60
Dbl	121 AAACCATCTGGGATGATGATACATCAATCAAGTTCAATAGAGATTTGAAGCACTTACTT 180
OY	61 GlnThrGlyLysSerProHiseSerGlnLeuLeuPheAspTrpGlyThrTrnaAspCysThr 80
Dbl	181 CAACATCGAAAAAGATCCCACTTCTAATTAAGTCTGGGGCACCACAAAATGACACA 240
OY	81 ValGlyAspLeuValAspLeuLeuIleGlnAnsgluPheAlaProIleSerLeuLeu 100
Dbl	241 GCCTGGATCTGTGGATCTTTTGTATCCAATATGAATTTTTGCTCCTGGAGCTTTTG 300
OY	101 LeuProAspAlaValProLysThrAlaAnthLeuProSerLysGluAlaIleThrVal 120
Dbl	301 CTCACGATGCTGTTCCCAAACCTGCTAATACACTTCTTAAGAAGCTAATACAGTT 360
OY	121 GlnGlnLysGlnMezProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140
Dbl	361 CAGCAAAAACAATGCTTTCTGTACAAACAAGAGATTTGAAGAACCACTTGGCAGAAAT 420
OY	141 LeuGlnLysSerTyMezProProAspSerSerSerProGluAnuLysSerLeuGluVal 160
Dbl	421 CTTGAAACAAGCTATATGCGACCTGACTCTCAAGTCCAGAAATATAAAGTTTGAAGATT 480
OY	161 SerAspThrArxPheHisSerPheSerPheTyrgIuleuLysAsnValThrAsnAspHe 180
Dbl	481 AGGATACACAGTTTTCACAGTTTTCATTTATATGAATTTGAAGAAATGTCACAAATTAACCTT 540
OY	181 AspGluArgProIleSerValGlyLysAnuLysMetGlyGluGluGlyPheGlyValVal 200
Dbl	541 GAATGAACGACCATTCTGTGGTGATATATAAATGGAGAGGAGATTTGGAGTTGTA 600

```

QY 201 TylrlysglytyrValasnaenthThrValaValaValyslyleuValaValaValasp 220
DB 601 TATTAAGGCTACGTAATAATTAACAACTGTGGCAGTAAGAAAGCTTGACGAAATGTTGAC 660
QY 221 lIethrThrglugluleuylsglInGlnPheaspGlnGluilelyValMetAlaIysCys 240
DB 661 ATTACTACTGAAAGACTGAAACACAGATTGTATCAGAAATAAAGTAATGGCAAGGT 720
QY 241 GlnhlsGluValenleuValGluLeuLeuGlyPheSerSerAspGlyAspAspLeuCysleu 260
DB 721 CAACATGAAACCTTAGTAAGAACTACTGTGTTCTCAAGTAGAGAGATGACCTCTGCTTA 780
QY 261 ValTyrValTyrMetProAsnGlySerleuLeuAspArgleuSerCysleuAspGlyThr 280
DB 781 GTATATGTTTACATGACCTTAATGTTTCATGCTAGACAGACTCTCTGTTGGATGATCT 840
QY 281 ProProleuSerTrpHisMetArgCysIysIleAlaGlnGlyAlaAlaAsnGlyIleAsn 300
DB 841 CCACCACTTCTTGGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 301 PheleuThsGluValenleuValGluLeuLeuGlyPheSerSerAspGlyValleuLeuAsp 320
DB 901 TTTCTACATGAAATCATCATATTCATAGATATTAATAATGCAAAATATCTTCTGAT 960
QY 321 GluAlaPheThrAlaIysIleSerAspPheGlyleuAlaArgAlaSerGluIysPheAla 340
DB 961 GAACCTTTTACTGTAATAATTCGACTTGTGCTTGACGGGCTTCGAGAAATTTGCC 1020
QY 341 GlnThrValMetThrSerArgIleValGlyThrThrAlaIysMetAlaProGluAlaLeu 360
DB 1021 CAGACAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 361 ArgGlyGluIleThrProIysSerAspIleTyrSerPheGlyValleuLeuGluIle 380
DB 1081 CGTAGAAGAAATTAACCCAAATCTGATATTTACAGCTTGTGTGTTTACTGAAATA 1140
QY 381 lIethrGlyleuProAlaValaIaspGluIysArgIleProGlnleuLeuAspIleIys 400
DB 1141 ATAAGTGGACTTCCAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 401 GluGluIleGluAspGluGluIysThrIleGluAspTyrIleAspIysMetAsp 420
DB 1201 GAAAGAAATTTGAAGATGAAGAAAGAAAGCAATTAAGATTATTTATTAAGATGAT 1260
QY 421 AlaAspSerThrSerValGluAlaMetThrSerValAlaSerGlnCysleuHlsGluIys 440
DB 1261 GCTGATTCACCTTCACTTGAAGCTATGATCTCTGTGCTAGTCAATGCTCATGAAAG 1320
QY 441 LysAsnIysArgProAspIleIysIleValGlnGlnleuLeuGlnIleMetThrAlaSer 460
DB 1321 AAAAATAAGAACCAACGATTAAGAGTTCAACAGCTGCTGCAAGAGATACAGCTTCT 1380

```

```

FT /*cag= a
FT /product= "Human IRAK4"
PN W0200240680-A2.
XX
XX 23-MAY-2002.
XX
XX 15-NOV-2001; 2001MO-US44844.
XX
XX 17-NOV-2000; 2000US-0715893.
PR 29-JUN-2001; 2001US-301889P.
XX
XX (BURN-) BURNHAM INST.
PA
XX Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
PI Stemmer-Jewen F;
XX
XX WPI; 2002-500222/53.
DR P-PSDB; AAE24865.
XX
XX New polypeptide comprising a death domain or death effector domain,
PT useful for discovery of drugs that suppress infection, inflammation,
PT allergy, sepsis, autoimmunity, allograft rejection and other diseases
PT
XX
XX Claim 19; Page 194-196; 209pp; English.
XX
XX The invention relates to an isolated polypeptide comprising a death
XX domain (DD), death effector domain (DED) or NB-ARC domain. The invention
XX is useful for identifying a binding agent, preferably a protein or a drug
XX that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
XX domain from DAP3, IRAK4, CTDD (Chlamydia trachomatis DD protein), DED4 or
XX NIDP (NGPR-interacting Death Domain), with a candidate binding agent and
XX detecting the association of the domain and the candidate binding agent,
XX by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
XX chemical crosslinking, nuclear magnetic resonance (NMR), mass
XX spectroscopy (MS) and FPA. The invention is useful for modulating the
XX level of a cell process such as cell proliferation, cell adhesion, cell
XX stress responses, responses to microbial infection and B cell
XX immunoglobulin class switching, in particular apoptosis within a cell.
XX Antibody specifically reactive with CTDD DD of C. trachomatis, C
XX muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the
XX CTDD DD protein is useful for detecting a Chlamydia infection. The
XX invention is useful for modulating the activity of oncogenic proteins,
XX for treating a pathology caused by the oncogenic proteins and for
XX treating bacterial infections by modulating the activity of bacterial
XX proteins. The protein and antibody specific for it are useful for
XX discovery of drugs that suppress infection, inflammation, allergy,
XX sepsis, autoimmunity, allograft rejection and other diseases. The protein
XX is useful for treating immune-based pathologies, pathologies associated
XX with cell division, inflammatory diseases such as sepsis, fibrosis,
XX arthritis, graft versus host disease. The invention is used in antisense
XX therapy and gene therapy. The present sequence is human IRAK4 gene.
XX
XX Sequence 2817 BP; 912 A; 547 C; 586 G; 772 T; 0 other;
XX
XX
XX Alignment Scores:
XX
XX Score: 5.07e-231 Length: 2817
XX Percent Similarity: 2362.00 Matches: 456
XX Best Local Similarity: 99.13% Conservative: 0
XX Query Match: 99.13% Mismatches: 4
XX DB: 24 Gaps: 0
XX
XX US-10-001-254-16 (1-460) x AAD40085 (1-2817)
QY 1 MetAsnIysProIleThrProSerThrTyrValArgCysleuAsnValGlyleuIleArg 20
DB 50 ATGACAAACCCATTAACATCACTATGTCGCTGCTCATATGTTGACCTAATTAAG 109
QY 21 LysleuSerAspPheIleAspProGlnGlnGlyTyrIysIysIleuAlaIleIys 40
DB 110 AAGCTGTCAGATTATTATGATCCTCAAGAGATGAGAAAGATTAGCTGATTAATAA 169

```

QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgArgPheGlnIleLeuLeu 60
 Db 170 AAACCATCTGGTGGATGATGATACATCATGTTTCAATAGAGAGATTGAAACATTAATCT 229
 QY 61 GlnThrGlyLysSerProThiSerGlyLeuLeuPheAspTyrGlyThrThraAsnCysThr 80
 Db 230 CAAACTGGAAAAAGTCCCATCTCTGAATTAATGCTGGGGCACCACAATTTGCACA 289
 QY 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPheAlaProIleSerLeuLeu 100
 Db 290 GTTGGTGATCTTGTGATCTTTTGTGATCCAAAATGAAATTTTGTCTCCGGAGTCTTTTG 349
 QY 101 LeuProAspAlaValProLysThrAlaAsnThrLeuProSerLysGlnAlaIleThrVal 120
 Db 350 CTCCAGATGCTGTTCCCAAACTGCTATACATCCTTCRAAGAGACTTAACAGT 409
 QY 121 GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140
 Db 410 CAGCAAAAAACAGATGCTTCTGTGCAAAAGACAGACATGATGACACTGTGCAGAAAT 469
 QY 141 LeuGlnGlnSerTyrMetProAspSerSerProGlnAsnLysSerLeuGlnVal 160
 Db 470 CTGGAACAAAGCTATATGACACCTGACTCTCAAGTCCGAAAAATPAAAGTTAGAACTT 529
 QY 161 SerAspThrArgPheHisSerPheSerPheTyrGlnLeuLysAsnValThraAsnAspHe 180
 Db 530 AGTGATACACGCTTTTCAAGTCTTTTCAATTTGATGAAGAATGTCACAAATTAATCTT 589
 QY 181 AspGluArgProIleSerValGlyLysAsnLysMetGlyGlnGlyPheGlyValVal 200
 Db 590 GATGAAGACGCCATTTCTGTGTGTGTATATAAAGGAGAGAGAGATTTGAGATGTGA 649
 QY 201 TyrLysGlyTyrValAsnAsnThrThrValAlaValLysLysLeuAlaIleMetValAsp 220
 Db 650 TATAAAGGCTAGTAAATPACAACTGTGCGAGGAAGAAACCTTGCAGCAATGGTTGAC 709
 QY 221 IleThrThrGlnGlnLysGlnGlnPheAspGlnGlnIleLysValMetAlaLysCys 240
 Db 710 ATTACTACTGAAGAACTGAAACAGACGTTTGTATCAAGAAATPAAAGTAAATGCAAGTGT 769
 QY 241 GlnHisGlnAsnLeuValGlnLeuLeuGlyPheSerSerSerAspGlyAspAspLeu 260
 Db 770 CAACATGAAGAACTTGTAGAACTACTTGTGTTCTCAAGTGAAGAGATGACCTTGCTTA 829
 QY 261 ValTyrValTyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThr 280
 Db 830 GTATATGTTTACATCTTAATGTTCAATGTTGCTAGACAGACTCTCTTGTGATGTACT 889
 QY 281 ProProLeuSerTyrPheIleMetArgCysLysIleAlaGlnGlnAlaAlaAsnGlyIleAsn 300
 Db 890 CCACCACTTTCTGGCAGCATGAGATGCAAGATGCTCAGGGTGCAGCTTAATGGCATCAAT 949
 QY 301 PheLeuHisGlnAsnHisIleHisIleArgAspIleLysSerAlaAsnIleLeuLeuAsp 320
 Db 950 TTTCTACATGAAGAAATCATCATATCATAGAGATTAATAAGTCAAAATATCTTACTGAT 1009
 QY 321 GlnAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGlnLysPheAla 340
 Db 1010 GAAGCTTTTACTGCTAAATATCATCTTGGCTTGCACGGGCTTTCAGAGAAAGTTTCC 1069
 QY 341 GlnThrValMetThrSerArgIleValGlyThrThrAlaTyrMetAlaProGlnAlaLeu 360
 Db 1070 CAGACAGTCATGACATGACAGATTTGGGACACACAGCTTATATGACCACAGACCTTTG 1129
 QY 361 ArgGlyGlnIleThrProLysSerAspIleTyrSerPheGlyValValLeuLeuGlnIle 380
 Db 1130 CGTGAGAAATPACACCCAAATCTGATATTTACACTTGGTGGTGGTTTACTTAAGATA 1189
 QY 381 IleThrGlyLeuProAlaValAspGlnHisArgGlnLupProGlnLeuLeuAspIleLys 400
 Db 1190 ATAACGTGACCTTCAGCTGTGTGATGAACACCGTGAACTCAGTATTTCTGATATTTAAA 1249

QY 401 GlnGlnIleGluAspGlnGlnLysThrIleGlnLysPyrIleAspLysLysMetAsnAsp 420
 Db 1250 GAAGAAATTTGAAGATGAAGAAAGAAAGCAATTTGAAGTTATATGATTAATAAAGATGATCAT 1309
 QY 421 AlaAspSerThrSerValGlnAlaMetCysSerValAlaSerGlnCysLeuHisGlnLys 440
 Db 1310 GCTGATTCACCTTCACTGATGAAGCTATGACTCTGTGTCTACCAATGTCCGACATGAAAA 1369
 QY 441 LysAsnLysArgProAspIleLysLysValGlnGlnLeuLeuGlnGlnMetThrAlaSer 460
 Db 1370 AAAAATTAAGACCCAGACATTAAAGAGTTTCCAGCTGTGCAAGATGACAGCTTCT 1429
 RESULT 4
 AAD10198
 ID AAD10198 standard; cDNA; 1542 BP.
 XX
 AC AAD10198;
 XX
 DT 24-SEP-2001 (first entry)
 XX
 DE Mouse interleukin-1 (IL-1) receptor-associated kinase (IRAK)-4 cDNA.
 XX
 KW Mouse; interleukin-1 receptor associated kinase-4; IRAK-4; cytosolic;
 KW IL; antibacterial; antiinflammatory; ophthalmological; vasotropic; OP;
 KW chronic obstructive pulmonary disease; neuroprotective; chronic cough;
 KW adult respiratory disease syndrome; pulmonary fibrosis; asthma; ARDS;
 KW interstitial lung disease; allergic rhinitis; transplant rejection;
 KW autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;
 KW multiple sclerosis; diabetes; cancer; solid tumour; lymphoma; stroke;
 KW cardiovascular disease; atherosclerosis; neurodegenerative disease;
 KW sepsis; osteoarthritis; rash; osteoporosis; psoriasis; dermatitis;
 KW inflammatory bowel disease; gout; Crohn's disease; ulcerative colitis;
 KW Behcet's syndrome; ankylosing spondylitis; IL-1 receptor/Toll receptor;
 KW sarcoidosis; transgenic animal; ss.
 XX
 OS Mus sp.
 XX
 FH Key
 FT CDS 1..1542 location/Qualifiers
 FT /tag= a
 FT /product= "Mouse IRAK-4"
 FT
 PD WO200151641-A1.
 XX
 PN 19-JUL-2001.
 XX
 PF 12-JAN-2001; 2001WO-US01171.
 XX
 PR 13-JAN-2000; 2000US-0176395.
 PA (TUDA-) TUDARIX INC.
 PI Weeche H, Li S;
 DR MPI; 2001-451860/48.
 DR P-PSDB; AAE05399.
 XX
 PT Novel human interleukin-1 receptor associated kinase polypeptide,
 PT useful for identifying modulators of the polypeptide for treating gout,
 PT asthma, allergic rhinitis, multiple sclerosis and skin cancer
 XX
 PS Claim 26; Fig 4; 89pp; English.
 CC The present sequence is a cDNA encoding mouse interleukin (IL)-1 receptor
 CC associated kinase (IRAK)-4. IRAK associate with activated IL-1, IL-18
 CC and other receptors and act to transduce signals originating from the
 CC activated receptors, ultimately leading to a variety of downstream
 CC effects such as nuclear factor (NF)-kappaB activation. The IRAK-4
 CC inhibitors are useful for treating inflammatory diseases such as
 CC pulmonary diseases and diseases of the airway (e.g., adult respiratory
 CC disease syndrome (ARDS), chronic obstructive pulmonary disease (OPD),
 CC pulmonary fibrosis, interstitial lung disease, asthma, chronic cough
 CC or allergic rhinitis), transplant rejection, autoimmune diseases (e.g.,

CC rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis
CC or diabetes), cancer (e.g., solid tumors, skin cancer or lymphoma),
CC cardiovascular diseases (e.g., stroke and atherosclerosis), diseases
CC of the central nervous system (e.g., neurodegenerative disease), CD14
CC mediated diseases, non-CD14 mediated sepsis, osteoarthritis, osteoporosis,
CC psoriasis, diseases of the skin (e.g., rash, contact dermatitis, atopic
CC dermatitis), inflammatory bowel disease (e.g., Crohn's disease and
CC ulcerative colitis), Behcet's syndrome, ankylosing spondylitis, gout,
CC sarcoidosis and ophthalmic diseases and conditions. The inhibitors of
CC IRAK-4 activity or expression are used to inhibit signal transduction
CC resulting from the activation of an interleukin-1 receptor (IL-1R)/Toll
CC receptor in a cell. They also inhibit the activation of a transcription
CC factor that activates NF-kappaB in the cell. IRAK-4 is used to create a
CC nonhuman transgenic animal which is useful for testing the function of
CC IRAK-4 in vivo, to generate models for the study of inflammatory
CC disorders and conditions and for the development of potential treatments
CC for IRAK-4-related inflammatory diseases and conditions. IRAK-4 sequences
CC are also used in gene therapy and in antisense therapy.

XX Sequence 1542 BP; 421 A; 392 C; 423 G; 306 T; 0 other;

Alignment Scores:

Pred. No.:	3,966-196	Length:	1542
Score:	2017.00	Matches:	385
Percent Similarity:	90.85%	Conservative:	32
Best Local Similarity:	83.88%	Mismatches:	42
Query Match:	84.53%	Indels:	0
DB:	22	Gaps:	0

US-10-001-254-16 (1-460) x AAD10198 (1-1542)

```
OY 1 MetAsnLYSProlleThrProSerThrTYrValAlArgCYLeuAsnValGlyLeuIleArg 20
Db 163 ATGAACAAGCCGTTGACACCATCGACATACATACCACTTAATGTGGGATCTTAGG 222
OY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPlySlyLeuAlaValAlaIleLys 40
Db 223 AAGCTGCGAGATTATATGATCTCTCAAGAAGGTGGAGAAATTAGCAGTAGCTATCAA 282
OY 41 LysProSerGlyYAspAspArgTyrAsnGlnPheHisIleArgAArgPheGluAlaIleu 60
Db 283 AAGCGCTCCGGCGAGCAGACATACATCAATGATTAAGAGAGATTGGAAGCCCTTACTT 342
OY 61 GlnThrGlyLysSerProThrSerGluLeuLeuPheAspTyrPGLYThrThrAsnCYthr 80
Db 343 CAGACCGGGAAGAGCCCACTGTAACCTGCTTTGACTGGGACCAACCAACTGCACCA 402
OY 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGluPheAlaProAlaSerLeuLeu 100
Db 403 GTTGGCGACTTGTGGATCTACTGCTCCAGATTGAGCTGTTGGCCCGCCACACTCTCTG 462
OY 101 LeuProAspAlaValProLysThrAlaAsnThrLeuProSerLysGluAlaIleThrVal 120
Db 463 CTGCCCGATGCCGTTCCCAAAACGCTCAAAAGCCCTGCTCTTGAAGAGCGCAACAGTG 522
OY 121 GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140
Db 523 GCACAAACAACACGGGCTTTCAGAGAAAAGACAGAGAACTCCGTAATGCTTATGCCGAG 582
OY 141 LeuGlnGlnSerTyrMetProProAspSerSerSerProGluAsnLysSerLeuVal 160
Db 583 CTAGAACACACGCTGCGAGCAACCGGACTCTTCAAGCCAGAACAGAAAGTGAAGTCC 642
OY 161 SerAspThrArgPheHisSerPheSerPheTyrGluLeuValAsnValThrAsnAspHe 180
Db 643 AGCAGCACTCGGTTCCACAGCTTCCTGCTTCAATGACAGAGAGAGCAAAACAACCTTC 702
OY 181 AspGluArgProIleSerValGlyLysAsnLysMetGlyGluGlyPheGlyValVal 200
Db 703 GACAGCAACCCGCTTCCGCGTGGCAACCGGATGGAGAGGGGGAATTGGAGTGGTG 762
OY 201 TyrLysGlyTyrValAsnAsnThrThrValAlaValLysLysLeuAlaIleMetValAsp 220
```

```
Db 763 TACAGGCGCTGTGTGACAAACACCATCTGGCCGTGAGAGAAAGCTCCGAGCGATGTTGAA 822
OY 221 IleThrThrGlnGluLeuLysGlnGlnPheAspGlnGluIleLysValMetAlaLysCys 240
Db 823 ATCAGTCTCTAAGAACTTAAGCAACAGTTTGAATCCAGAAATTAAGTAAGGCAACCTGT 882
OY 241 GlnHisGluAsnLeuValGluLeuLeuGlyPheSerSerAspGlyAspAspLeuCysLeu 260
Db 883 CAGACACAGAACTGGGTGGAGCTGCTCGGCTTCCAGCGACGACGCAACCTGTGCTTA 942
OY 261 ValTyrValLysMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThr 280
Db 943 GTGTAGCTTACATGCCCCAAGCGGCTCTTGTGTGACAGACTGTCTGCTGGATGGTACA 1002
OY 281 ProProLeuSerTrpHisMetArgCysLysIleAlaGlnGlyAlaIleAsnGlyIleAsn 300
Db 1003 CCACCGCTTCTCTGGACACAAAGGTGAAGTGTCTCAGGAGCAAGAAATGGCATCAGG 1062
OY 301 PheLeuHisGluAsnHisIleHisIleArgAspIleLysSerAlaAsnIleLeuLeuAsp 320
Db 1063 TTCTCGATGAAATCATCATCATCATATTAAGTCAAAATATCTTACTAGAC 1122
OY 321 GluAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGluLysPheAla 340
Db 1123 AAGACTTCTACTGCAAAATATCTGACTTTGGCTTGGCAGGGCTTCCGCAAGGCTTAGC 1182
OY 341 GlnThrValMetThrSerArgIleValGlyTyrThrAlaTyrMetAlaProGluAlaLeu 360
Db 1183 CAGACGCTCATGACACACCGCAATCTGTGGCAACCGCTTTCATGGAACCCGAAGCTTTG 1242
OY 361 ArgGlyGluIleThrProLysSerAspIleTyrSerPheGlyValValLeuLeuGluIle 380
Db 1243 CGGGGAATAATTAACACCAATCTGACATCTACAGCTTCGGCGTGTCTGTTGGAGCTG 1302
OY 381 IleThrGlyLeuProAlaValAspGlnHisIleArgGluProGlnLeuLeuAspIleLys 400
Db 1303 ATTAACCGGCTGGCGGCTGTGATGAAGAACCGTGAACCTGAACTACTGCTGATATTA 1362
OY 401 GlnGluIleGluAspGlnGluLysThrIleGluAspTyrIleAspLysLysMetAsnAsp 420
Db 1363 GAGAGATTAAAGTGAAGAGAAAGACATTTGAAGATTTACACGATGAGAGATGACGAT 1422
OY 421 AlaAspSerThrSerValGluAlaMetTyrSerValAlaSerGlnCysLeuHisGluLys 440
Db 1423 GCGGACCTGCTCGGGAAGCAATGTACTGTGCTGCTAGCAAGTGTGCAATGAGAAAG 1482
OY 441 LysAsnLysArgProAspIleLysLysValGlnGlnLeuLeuGlnGlnMetThrAla 459
Db 1483 AAAAACAAGCGGCCAGACATTTGCAAGGTTCAACAGCTGTACAGAGATGTCTGCT 1539
OY 459
```

RESULT 5
AAH13798
ID AAH13798 standard; cDNA; 2213 BP.
AC AAH13798;
XX 26-JUN-2001 (first entry)
DE Human cDNA sequence SEQ ID NO:10742.
DT 26-JUN-2001 (first entry)
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
OS Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
PR

PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX Oca T, Iocgai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI: 2001-118749/34.
 DR
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 10742; 2537bp + CD ROM; English.
 XX
 XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 2213 BP; 642 A; 488 C; 516 G; 567 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,88e-172 Length: 2213
 Score: 1787.50 Matches: 352
 Percent Similarity: 87.34% Conservative: 0
 Best Local Similarity: 87.34% Mismatches: 1
 Query Match: 74.92% Indels: 50
 DB: 22 Gaps: 1
 US-10-001-254-16 (1-460) x AAH13798 (1-2213)

DB 913 -----TGCTGTCCCAAAACGCTAATACACTTACTTAAAGAGCATTAACGT 963
 QY 1GlnGlnGlnSerMetProPheCysAspIleAspArgThrLeuMetThrProValGlnAs 140
 DB 964 TCAGCAAAAAGAGAGGCTTTCTGTGACAAAGACGAGCATTTGAGACACCTGTGCAGAA 1023
 QY 140 nLeuGlnGlnSerThrMetProProAspSerSerSerProGluAsnIleYSerLeuGluVa 160
 DB 1024 TCTTAAACAAAGCTATATGCCACTGACTCTCTCAAGTCCAGAAATTAAGTTTGAAGT 1083
 QY 160 lSerAspThrArgPheHisSerPheSerPheThrGluLeuYAsnValThrAsnAspH 180
 DB 1084 TAGGATACACGTTTTCACAGTTTTCATTTTATGAAATGGAAGATGACAAATTAACCT 1143
 QY 180 eAspGluArgProIleSerValGlyGlyAsnIleMetGlyGlyGlyPheGlyValVa 200
 DB 1144 TGATGAAGACCCATTTCTGTGTGTGTATTAATAATGGAGAGGAGATTGGAGTTGT 1203
 QY 200 lTyrIleGlyTyrValAsnAsnThrThrValAlaValIleYsIleuAlaIleMetValAs 220
 DB 1204 ATATTAAGGCTTACGTAATTAACACACTGTGCGAGTGAGAAAGCTTGCAAGATGTGGA 1263
 QY 220 pIleThrThrGluGluLeuYsGlnGlnPheAspGlnIleuYsValMetAlaYsCy 240
 DB 1264 CATTACTACTGAGAGACTGAAACAGCACTTTGATCAAAATTAAGTAATGGCAAAATG 1323
 QY 240 eGlnHisGluAsnLeuValGluLeuGlyPheSerSerAspGlyAspAspLeuCyLe 260
 DB 1324 TCACATCAATAAACTTAGTAGAATCACTGTTCTTCAAGTATGAGATGACCTTCTGCTT 1383
 QY 260 vValIleValIleThrProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyTh 280
 DB 1384 AGTATATTTTACATGCTTAACGTTCACTTCTAGACAGACTCTTCTGTGAGATGATC 1443
 QY 280 rProProLeuSerThrPheMetArgCysIleAlaGlnGlyAlaIleAsnGlyIleAs 300
 DB 1444 TCACACACTTTCTTGCCACATGAGATGAGATGCAATGCTGAGGCTGACCTAATGCAATCA 1503
 QY 300 nPheLeuHisGluAsnHisIleHisIleIleAspIleYsSerAlaAsnIleLeuLeuAs 320
 DB 1504 TTTTCTACATGAATAATCATCATATTCATATGAGATATTAAGTGCATATCTTACTGCA 1563
 QY 320 pGluAlaIlePheThrAlaYsIleSerAspPheGlyLeuAlaArgAlaSerGlyYsPheAl 340
 DB 1564 TGAAGCTTTTACTGCTAAATATCTGACTTTTGCCCTTGCAAGGCTTCTGAGAGATTGC 1623
 QY 340 aglIleThrValMetThrSerArgIleValGlyThrThrAlaThrMetAlaProGluAlaLe 360
 DB 1624 CCAGACAGCTCATGACTAGCAAGATTTGTGGAAACAACAGCTTATATGCAACAGAACTTT 1683
 QY 360 uArgGlyGlyIleThrProYsSerAspIleTyrSerPheGlyValValLeuLeuGluI 380
 DB 1684 GCGTGAGAAATTAACCCCAATCTGATATTAACGCTTGTGTGTGTTACTAGAAAT 1743
 QY 380 eIleThrGlyLeuProAlaValAspGluHisArgGluProGlnIleuLeuAspIleY 400
 DB 1744 AATACTGAGACTTCCAGCTGTGATGAACACCGTAACCTGATTAATGCTAGATATTA 1803
 QY 400 sGluGlu 402
 DB 1804 AGAAGA 1810
 RESULT 6
 AAS76805
 ID AAS76805 standard; cDNA; 1668 BP.
 AC AAS76805;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #12609.
 XX

RESULT 7
ID AAA09319 standard; DNA; 833 BP.
XX
AC AAA09319;
XX
DT 10-AUG-2000 (first entry)
XX
DE Human cancer associated antigen precursor DNA, clone NY-REN-64.
XX
KW renal cancer; cancer associated antigen precursor; diagnosis;
XX cytostatic; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 50..670
FT /tag= a
XX
PN WO200020587-A2.
XX
PD 13-APR-2000.
XX
PF 04-OCT-1999; 99WO-US22873.
XX
PR 05-OCT-1998; 98US-0166300.
PR 05-OCT-1998; 98US-0166350.
XX
XX (LUDWIG INST CANCER RES.
XX
PI Obata Y, Gout I, Tureci O, Sahin U, Pfeundschuh M, Scanlan MJ;
PI Stockert E, Chen Y, Old LJ, Jager B, Knuth A;
DR WPI: 2000-303774/26.
DR P-PSDB; AAY92347.
XX
PT Preventing, diagnosing and/or treating disorders associated with
PT abnormal expression of human cancer associated antigens
XX
PS Claim 57; Page 85; 121pp; English.
XX
CC AAA09310-20 are novel genes isolated by SEREX screening from a renal
CC cancer cell line 1973/10.4. The genes encode cancer associated antigen
CC precursors. These gene products are useful in methods for preventing,
CC diagnosing and/or treating disorders, especially cancer, associated with
CC abnormal expression of human cancer associated antigens. The method
CC comprises contacting a sample from a subject with an agent that
CC specifically binds to the nucleic acid molecule or expression product
CC (or fragment) complexed with a human leukocyte antigen (HLA) molecule
CC and determining the interaction between the agent and the nucleic acid
CC molecule or the expression product as a determination of the disorder.
XX
SQ Sequence 833 BP; 273 A; 155 C; 179 G; 226 T; 0 other;
XX
Alignment Scores:
Pred. No.: 7,786-110 Length: 833
Score: 1169.00 Matches: 246
Percent Similarity: 93.21% Conservative: 1
Best Local Similarity: 92.83% Mismatches: 8
Query Match: 48.99% Indels: 10
DB: 21 Gaps: 3
US-10-001-254-16 (1-460) x AAA09319 (1-833)
QY 1 MetAaLysPProIleThrProSerThrTyrValArgCysLeuAsnValGlyLeuIleArg 20
DB 50 ATGAACAAACCATTAACCATCAATATGTGCGCTGCTCAATGTGGAATAATTAGG 109
QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPheValLeuAlaValAlaIleLys 40
DB 110 AAGCTGTGATTTATTTGATCTCTCAAGAGATGAGAAAGTATGCTGTAGCTTTTAAA 169
QY 41 LysProSerGlyAspAspArgTyrAsnGlnIlePheHisIleArgArgPheGluAlaLeuLeu 60

DB 170 AAACATCTGGATGATAGATAGATACATCATGTTTACATAGAGATTTGAACATTACTT 229
QY 61 GlnThrGlyLysSerProThrSerGluLeuLeuPheAspTyrGlyThrAsnGlyThr 80
DB 230 CAAACTGGAAGAAAGTCCCACTTGTGAATTAAGTGTGAGCTGGGGCCACCAATATGACA 289
QY 81 ValGlyAspLeuValAspLeuLeuIleGlnIleAsnGlnPheAlaProAlaSerLeuLeu 100
DB 290 GTTGGTGAATCTTTGTGATCTTTGATCCAAATGAATTTTGTCTCTGGCAGTCTTTTG 349
QY 101 LeuProAspAlaValProLysThrAlaAsnThrLeuProSerLysGluAlaIleThrVal 120
DB 350 CTCACGATGCTGTTCCCAAACTGCTAATACATACCTTTAAAGAACTATACAGTT 409
QY 121 GlnGlnLysGlnMetProPheCysAspLysAspArgTyrLeuMetThrProValGlnAsn 140
DB 410 CAGCAAAAACAGATGCTTTCTGTGACAAAGACAGACATTTGATGACACTGTGACGAT 469
QY 141 LeuGlnGlnSerTyrMetProProAspSerSerSerProGlnIleAsnLysSerLeuGluVal 160
DB 470 CTGGAACAAAGCTATATGCTCCACTGCTCTCAAGTCCAGAAATATTAAGTTAGAGTT 529
QY 161 SerAspThrArgPheHisSerPheSerPheTyrGluLeuLysAsnValThrAsnAspHe 180
DB 530 AGTATACACCGTTTTCACAGTTTTCATTTATGAATTTGAGAAATGTCACAAATACCTT 589
QY 181 AspGluAspProLysSerValGlyLysAsnLysMetGlyGluGlyPheGlyValVal 200
DB 590 GATGAACACACCATTTCTGTGTGTGTATTAATTAAGGAGAGGAGATTTGTGAGTTCTTA 649
QY 201 Tyr-LysGlyTyrValAsnAsnThrThrValAlaVal-LysLysLeuAlaIleMetVal 220
DB 650 TATTAAGGCTTACGTAAATTAACACACTGTGCACTGAAAGAGCTTCCAGCAATGTTG 709
QY 220 spIleThrThrGlu-GluLeuLys-GlnGlnPheAsp-GlnGlu-IleLysValMetAla 238
DB 710 ACATTACTACTGAAGAACTGAAACACGACAGTTGATCCAGAAATTAAGTATGAGGC 769
QY 239 LysCysGlnHis-----GluAsnLeuVal---GluLeuLeuGlyPheSerSer---Asp 254
DB 770 AAAAGTGTCAACCATGGAAGAAACCTTAGTAGGAACCTACTGTGGTTCCTCAAGTGAT 829
QY 255 Gly 255
DB 830 GGG 832
RESULT 8
ID AAS27214/C
AAS27214 standard; cDNA; 1493 BP.
XX
AC AAS27214;
XX
DT 07-NOV-2001 (first entry)
XX
DE cDNA encoding novel signal transduction pathway protein, Seq ID 249.
XX
XX
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
XX antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
XX immune system disorder; rheumatoid arthritis; inflammatory condition;
XX organ transplant rejection; infection; hepatitis C; blood disorder;
XX sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
XX reproductive system; gastrointestinal; liver disorder; AIDS; ss;
XX acquired immune deficiency syndrome.
OS Homo sapiens.
XX
XX WO2000154733-A1.
PN
XX
XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01312.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226661.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234222.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235835.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241321.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 09-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465460/50.
DR P-PSDB; AAU17297.

XX Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders
XX
PS Claim 1: SEQ ID No 249; 880bp; English.

XX The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
CC pathway protein coding sequences and PCR primers of the invention.
XX

Alignment Scores:

Pred. No.: 2, 4e-89 Length: 1493
Score: 972.00 Matches: 190
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 40.74% Indels: 0
DB: 22 Gaps: 0

US-10-001-254-16 (1-460) x AAS27214 (1-1493)

QY 271 LeuAaPaTgLeuSeRcYsleuAaPglYThrProPoleuSeRTrpHisMetArgCysIys 290
Db 1491 CTAGACACACTCTCTTGGTGGATGCTTCCACACCTTCTTGGACATGATGACAG 1432
QY 291 IleAlaGInGlyAlaAlaAsnGlyIleAsnPhelMetHisGluAsnHisIleHisArg 310
Db 1431 ATTGCTCAGGGGTCAGCTAATGACATCAATTTTCAATGAAATCATCATTTATGAGA 1372
QY 311 AspIleYsSerAlaAsnIleleuAaPgluaIaphetThAlaIleIleSerAspPhe 330
Db 1371 GATATTAAAGTGCATATCTTACTGATGATGAGCTTTTACGCTAAATATCTGACTTT 1312
QY 331 GlyLeuAlaArgAlaSerGluYsPheAlaGInThrValMetThSerArgIleValGly 350
Db 1311 GGCCCTTGACGGGCTTGGAGAGTTGCCACAGACATCATAGCAGAAATTTGGGA 1252
QY 351 ThrThraIaTyrrMetAlaProGluAlaLeuArgGlyGluIleThrProYsSerAspIle 370
Db 1251 ACAACAGCTTAATATGACACCAAGACCTTGCCTGAGAAATTAACCCCAATCTGATATT 1192
QY 371 TyrSerPheGlyValIleleuAaPgluIleIleThrGlyLeuProAlaValAspGluHis 390
Db 1191 TACAGCTTTGGTGCTTTTACTAGAAATTAATACTGAGCTTCCAGCTGTGATTAAC 1132
QY 391 ArgGluProGInleuAaPgluIleYsGluGluIleGluAspGluGluYsThrIle 410
Db 1131 CGTGAACCTCACTTCTGATGATATTAAAGAAATTAAGAGAGAGAAACCAAT 1072
QY 411 GluAspTyrrIleAspIleYsMetAsnAspAlaAspSerThrSerValGluAlaMetYr 430
Db 1071 GAAGATTATATTGATTAATAAAGTGAATGATGCTGATTCACCTTCAGTTGAAGCTATGATAC 1012

QY 431 SerValAlaSerGIncYsleuHisGluYsAsnIYsArgProAspIleYsIysVal 450
Db 1011 TCTGTTGCTAGTCATGATGCTGCATGATAAAGAAATTAAGAGCCAGACATTAAGAGTT 952
QY 451 GInGInleuAaPgluGInMetThAlaSer 460
Db 951 CAACAGCTGCTGCAGAGATGACAGCTTCT 922
RESULT 9
ID ABR44003/C
ABR44003 standard; cDNA, 1493 BP.
XX ABR44003;
AC
XX
DT 05-JUN-2002 (first entry)
XX
DE DNA encoding novel central nervous system protein #583.
XX
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminization;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy; gene; ss.
OS Homo sapiens.
PN WO200155318-A2.
XX
XX 02-AUG-2001.
PF 17-JAN-2001; 2001WO-US01332.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217486.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225216.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225475.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.

US-10-001-254-16 (1-460) X ABK44003 (1-1493)

OY	271	LeuAPARgLeuSerCysValuAspGlyThrProProLeuSerThrHisMetArgCysIys	290
Db	1491	CTAGACAGACTCTCTTGGCTGGATGGTACTCCACCACTTTCTTGGACATGAGATCAAG	1433
OY	291	lleaIaGInGlyAlaAlaAsnGlylIeaenPheLeuHISgluAsnHIShISlIeHISarg	310
Db	1431	ATTGCTCAGGGGCGAGCGTAAATGGCAATCAATTTTCTTCATGAAATCATCATATTCATAGA	137
OY	311	AspIleIysSerSerAlaAsnIleLeuLeuAspGluAlaPheThrAlaIysIleSerApphe	330
Db	1371	GATATTAAAGGCAAAATATCTTACTGATGAAAGCTTTTACGCTTAAATATCTGACTTT	1311
OY	331	GlyLeuAlaArgAlaSerGluIysPheAlaGlnThrValMetThrSerArgIleValGly	350
Db	1311	GGCCTTGACGGGCTTCTGAGAAAGTTTGGCCACAGCATCATGATGACGAATTTGGGA	125
OY	351	ThrThrAlaTyrMetAlaProGluAlaLeuArgGlyGluIleThrProIysSerAspIle	370
Db	1251	ACAAACAGCTTAATATGGCACCAAAAGCTTTGGCTGGAGAAATTAACCCAAATCTGATATT	119
OY	371	TyrSerPheGlyValValLeuLeuGluIuIeIleThrGlyLeuProAlaValAspGluHIS	390
Db	1131	CTGTAACCTCACTATTGGCTAGATTTTAAAGAAATTTGAAGATGAAGAAAGCAAT	107
OY	411	GluAspTyrIleAspIysIleMetAsnAspAlaAspSerThrSerValGluAlaMetCyr	430
Db	1071	GAAGATTATATGAAATTAAGATGAATGATGCTGATTCACCTTCAGTTGAACTTATGTC	101
OY	431	SerValAlaSerGlnCysLeuHISgluIysIysAsnIysArgProAspIleIysIysVal	450
Db	1011	TCTGTTCTAGTCATATGCTGATGATAAAGAAATTAAGAGCCAGACATTAAAGAGTT	952
OY	451	GlnGlnLeuLeuGlnGlnIuMetThrAlaSer	460
Db	951	CAACAGCTGTCAGAGATGACAGCTTCT	922
RESULT 10			
ABK43727			
ID	ABK43727	standard; cDNA; 1618 BP.	
XX	ABK43727;		
AC			
XX			
DT	05-JUN-2002	(first entry)	
DE			
XX		DNA encoding novel central nervous system protein #307.	
KM		Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;	
KM		hyperproliferative disorder; neoplasm; cardiovascular disorder;	
KM		cardiac arrest; cerebrovascular disorder; ischemia; angiogenesis;	
KM		nervous system disorder. Alzheimer's disease; AIDS; ocular disorder;	
KM		acquired immunodeficiency virus; dysplasia; gastrointestinal disorder;	
KM		adenocarcinoma; reproductive system disorder; testicular feminization;	
KM		endocrine disorder; diabetes; cancer; leukæmia; neovascularisation;	
KM		respiratory disorder; renal disorder; kidney failure; blood disorder;	
KM		myocardial infarction; wound healing; cell proliferation; skin aging;	
KM		food additive; food preservative; gene therapy; gene; ss.	
XX			
OS		Homo sapiens.	
PN			
XX	WO200155318-A2.		
PD	02-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US01332.		
XX			
XX	31-JAN-2000; 2000US-0179065.		
RR	04-FEB-2000; 2000US-0180628.		

PR	24-FEB-2000	2000US-0184664	PR	24-FEB-2000	2000US-0184664
PR	02-MAR-2000	2000US-0186350	PR	02-MAR-2000	2000US-0186350
PR	16-MAR-2000	2000US-0189874	PR	16-MAR-2000	2000US-0189874
PR	17-MAR-2000	2000US-0190076	PR	17-MAR-2000	2000US-0190076
PR	18-APR-2000	2000US-0198123	PR	18-APR-2000	2000US-0198123
PR	19-MAY-2000	2000US-0205515	PR	19-MAY-2000	2000US-0205515
PR	07-JUN-2000	2000US-0209467	PR	07-JUN-2000	2000US-0209467
PR	28-JUN-2000	2000US-0214886	PR	28-JUN-2000	2000US-0214886
PR	30-JUN-2000	2000US-0215135	PR	30-JUN-2000	2000US-0215135
PR	07-JUL-2000	2000US-0216480	PR	07-JUL-2000	2000US-0216480
PR	11-JUL-2000	2000US-0217487	PR	11-JUL-2000	2000US-0217487
PR	14-AUG-2000	2000US-0225214	PR	14-AUG-2000	2000US-0225214
PR	14-AUG-2000	2000US-0225213	PR	14-AUG-2000	2000US-0225213
PR	14-AUG-2000	2000US-0225266	PR	14-AUG-2000	2000US-0225266
PR	14-AUG-2000	2000US-0225667	PR	14-AUG-2000	2000US-0225667
PR	14-AUG-2000	2000US-0225658	PR	14-AUG-2000	2000US-0225658
PR	14-AUG-2000	2000US-0225477	PR	14-AUG-2000	2000US-0225477
PR	14-AUG-2000	2000US-0225470	PR	14-AUG-2000	2000US-0225470
PR	14-AUG-2000	2000US-0225757	PR	14-AUG-2000	2000US-0225757
PR	14-AUG-2000	2000US-0225758	PR	14-AUG-2000	2000US-0225758
PR	14-AUG-2000	2000US-0225759	PR	14-AUG-2000	2000US-0225759
PR	18-ANG-2000	2000US-0226679	PR	18-ANG-2000	2000US-0226679
PR	22-ANG-2000	2000US-0226681	PR	22-ANG-2000	2000US-0226681
PR	22-ANG-2000	2000US-0226688	PR	22-ANG-2000	2000US-0226688
PR	23-ANG-2000	2000US-0227182	PR	23-ANG-2000	2000US-0227182
PR	30-ANG-2000	2000US-0229824	PR	30-ANG-2000	2000US-0229824
PR	01-SEP-2000	2000US-0229287	PR	01-SEP-2000	2000US-0229287
PR	01-SEP-2000	2000US-0229343	PR	01-SEP-2000	2000US-0229343
PR	01-SEP-2000	2000US-0229344	PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345	PR	01-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229509	PR	05-SEP-2000	2000US-0229509
PR	05-SEP-2000	2000US-0229613	PR	05-SEP-2000	2000US-0229613
PR	06-SEP-2000	2000US-0230437	PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230438	PR	06-SEP-2000	2000US-0230438
PR	08-SEP-2000	2000US-0231142	PR	08-SEP-2000	2000US-0231142
PR	08-SEP-2000	2000US-0231143	PR	08-SEP-2000	2000US-0231143
PR	08-SEP-2000	2000US-0231144	PR	08-SEP-2000	2000US-0231144
PR	08-SEP-2000	2000US-0231145	PR	08-SEP-2000	2000US-0231145
PR	08-SEP-2000	2000US-0231146	PR	08-SEP-2000	2000US-0231146
PR	08-SEP-2000	2000US-0231147	PR	08-SEP-2000	2000US-0231147
PR	08-SEP-2000	2000US-0231148	PR	08-SEP-2000	2000US-0231148
PR	14-SEP-2000	2000US-0232399	PR	14-SEP-2000	2000US-0232399
PR	14-SEP-2000	2000US-0232399	PR	14-SEP-2000	2000US-0232399
PR	14-SEP-2000	2000US-0232400	PR	14-SEP-2000	2000US-0232400
PR	14-SEP-2000	2000US-0232401	PR	14-SEP-2000	2000US-0232401
PR	14-SEP-2000	2000US-0233063	PR	14-SEP-2000	2000US-0233063
PR	14-SEP-2000	2000US-0233064	PR	14-SEP-2000	2000US-0233064
PR	14-SEP-2000	2000US-0233065	PR	14-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234223	PR	21-SEP-2000	2000US-0234223
PR	21-SEP-2000	2000US-0234224	PR	21-SEP-2000	2000US-0234224
PR	21-SEP-2000	2000US-0234225	PR	21-SEP-2000	2000US-0234225
PR	21				

PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240962.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-581633/65.
DR P-PSDB; AAU87397.
XX
PT New isolated nucleic acid encoding a protein for diagnosing,
PT preventing, treating or ameliorating medical conditions and used as
PT food additives or preservatives -
XX

PS Claim 1; SEQ ID No 317; 837bp; English.
XX
XX The invention describes an isolated nucleic acid molecule (I) encoding a
CC novel central nervous system protein. (II) and polypeptides (III) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. Alzheimer's disease and
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired Immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,

Alignment Scores:

Pred. No.: 2,698-89 Length: 1618
Score: 972.00 Matches: 190
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 40.74% Indels: 0
DB: 23 Gaps: 0

US-10-001-254-16 (1-460) x ABK43727 (1-1618)

QY 271 LeuAspArgLeuSerCysLeuAspGlyThrProProLeuSerTrpHisMetArgCysLys 290
Db 3 CTGACACAGACTCTCTCTGATGTGATCCACCACTTCTTGACATGATGACAG 62
QY 291 TLeAGlnGlyAlaAlaAsnGlyIleAsnPhleuHisGluLysnHisIleHisArg 310
Db 63 ATTGCTCAGGGTGCAGCTAATGACATCAATTTCTCATATAAATCATCATTCATGAGA 122
QY 311 AspIleYSerAlaAsnIleLeuLeuAspGluAlaPheThrAlaLysIleSerAspPhe 330
Db 123 GATTTAAAGTGCATTAATCTTACTGATGATGAAGCTTTTACTGTAATAATCTGACTTT 182
QY 331 GlyLeuAlaArgAlaSerGluLysPheAlaGlnThrValMetThrSerArgIleValGly 350
Db 183 GGCCTTCACGGGCTTCTGAGAGTTGCCACACAGCTCATGACTGACAAATTGTGGGA 242
QY 351 ThrThrAlaYrMetAlaProGluAlaLeuArgGlyGluIleThrProLysSerAspIle 370
Db 243 ACAACACCTTAATGACACACAGAGCTTTCGGGACAAATTAACCCCAATCTGATATT 302
QY 371 TySerPheGlyValIleLeuLeuGluIleIleThrGlyLeuProAlaValAspGluHis 390
Db 303 TACAGCTTTGCTGTGGTTTACTAGAAATTAATGACTTCACGCTGTGATGAACAC 362
QY 391 ArgGluProGlnLeuLeuAspIleLysGluGluIleGluAspGluGluLysThrIle 410
Db 363 CGTGAACCTCAGTTATGCTAGATATTAAAGAAATTAAGATGAAGAAAGACAAATT 422
QY 411 GluAspTyrIleAspLysLysMetAsnAspAlaAspSerThrSerValGluAlaMetTyr 430
Db 423 GAAGATTATATTGATTAATAAAGAAAGATGATGATGATCCACTTCAGTTGAAGCTATGTAC 482
QY 431 SerValAlaSerGlnCysLeuHisGluLysLysAsnLysArgProAspIleLysVal 450
Db 483 TCTGTTCTATGTCATATCTGCATGAAGAAATAAGAACACGACATTAAGAAAGTT 542
QY 451 GlnGlnLeuLeuGlnGluMetThrAlaSer 460

Db 543 CAAACGCTGCTGCAAGATGACAGCTTCT 572

RESULT 11

ABL90822

ABL90822 standard; cDNA; 1181 BP.

AC ABL90822;

DE 24-MAY-2002 (first entry)

XX Human polynucleotide SEQ ID NO 1384.

XX Cytostatic; immunosuppressive; nocotropic; neuroprotective; antiviral;
XX anti-allergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
XX vulnerability; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein; gene; ss.

OS Homo sapiens.

PN WO200190304-A2.

XX 29-NOV-2001.

PF 18-MAY-2001; 2001WO-US16450.

PR 19-MAY-2000; 2000US-205515P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;

XX WPI; 2002-122018/16.

DR P-PSDB; ABB90413.

PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders -

XX Claim 4; SEQ ID NO 1384; 2081pp + Sequence Listing; English.

XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
XX (ABB89040-ABB90444) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune
XX haemolytic anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemia;
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX and parasitic infections.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1181 BP; 396 A; 222 C; 212 G; 348 T; 3 other;

XX Alignment Scores:

XX Pred. No.: 3,68e-88

XX Score: 959.00

XX Percent Similarity: 98.95%

XX Best Local Similarity: 98.95%

XX Query Match: 40.19%

XX DB: 24

XX Length: 1181

XX Matches: 188

XX Conservative: 0

XX Mismatches: 2

XX Indels: 0

XX Gaps: 0

US-10-001-254-16 (1-460) x ABL90822 (1-1181)

QY 271 LeuAspArgLeuSerCysLeuAspGlyThrProProLeuSerTrpHisMetArgCysIys 290
Db 3 CTAGACAGACTCTCTGCTTGATGATGACTCCACACTTTCTTGACACATGAGATGCAG 62
QY 291 ILeaIaGInGlyAlaAlaAsnGlyIleAsnPhleuHisGluAsnHisIleHisArg 310
Db 63 ATTGCTCAGGGGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 122
QY 311 AspIleLeuSerAlaAsnIleLeuLeuAspGlyAlaPheThrAlaIleSerAspPhe 330
Db 123 GATATTAAAGTCAAAATATCTTACTGATGATGATGATGATGATGATGATGATGATGATGAT 182
QY 331 GlyLeuAlaArgAlaSerGlyIleAsnPhleuHisGluAsnHisIleHisArg 350
Db 183 GGCCTTGACGGGCTTCTGASAGATTTCACAGACAGTATGATGATGATGATGATGATGATGAT 242
QY 351 ThrThrAlaIleAlaProGlyAlaLeuArgGlyIleThrProLeuSerAspIle 370
Db 243 ACAACAGCTTATATGACACCAAGACTTGGGTGAGAAATTAACCCCAATCTGATATT 302
QY 371 TyrSerPheGlyValIleLeuLeuGluIleIleThrGlyLeuProAlaValAspGluHis 390
Db 303 TACAGCTTGGTGTGTTTACTGAAATTAATTAATGACTTCCAGCTGTGATGATGATGATGAT 362
QY 391 ArgGluProGlnLeuLeuAspIleLeuGluGluIleGluAspGluGluIleThrIle 410
Db 363 CGTGAACCTCAGATTATCTGATGATTAAGAGAAATTAAGAGAAATTAAGAGAAATTAAGAG 422
QY 411 GluAspIleIleAspIleLeuAsnAspAlaAspSerThrSerValGluAlaMetTyr 430
Db 423 GAAGATTATATGATTAATAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 482
QY 431 SerValAlaSerGlnCysLeuHisGluIleAsnIleAsnIleAsnIleAsnIleVal 450
Db 483 TCTGTGCTAGTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 542
QY 451 GlnGlnLeuGlnGlnGlnMetThrAlaSer 460
Db 543 CAAACGCTGCTGCAAGATGACAGCTTCT 572

RESULT 12

AAF22618

ID AAF22618 standard; cDNA; 539 BP.

XX AAF22618;

XX 26-MAR-2001 (first entry)

XX Human breast cancer associated antigen nucleotide sequence SEQ ID NO:197.

XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
XX cancer associated antigen; cytosolic; cancer vaccine; ss.

XX Homo sapiens.

XX WO200073801-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US14749.

XX 28-MAY-1999; 99US-0136526.

XX 10-SEP-1999; 99US-0153454.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Obata Y;

XX WPI; 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated
XX antigen precursors, useful for diagnosing and treating a condition

PT characterized by expression of an abnormal amount of a protein, e.g.
 cancer -
 XX
 PS Claim 50; Page 338; 799pp; English.

XX Human, breast cancer; gastric cancer; prostate cancer; diagnosis
 KW cancer associated antigen; cytostatic; cancer vaccine; ss.
 XX

CC AAF22422 to AAF22656, AAF22657 to AAF22773 and AAF22774 to AAF33014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAF63332 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer.

XX WO200073801-A2.
PN
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US14749.
XX
PR 28-MAY-1999; 99US-0136526.
PR 10-SEP-1999; 99US-0153454.
XX
/XREF \ TYPING INNOVATION DPC

SQ Sequence 539 BP; 168 A; 94 C; 111 G; 163 T; 3 other;

Assignment Scores:	
Pred. No.:	7.17e-81
Score:	883.00
Percent Similarity:	97.24%
Best Local Similarity:	97.24%
Query Match:	37.01%
DB:	22
Length:	538
Matches:	176
Conservative:	0
Mismatches:	3
Indels:	2
Gaps:	0

US-10-001-254-16 (1-460) X AAF22618 (1-539)

PS Claim 50; Page 303; 799pp; English.

258 LeuCysLeuValTyrValTyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeu 277

CC AAF22422 to AAF22626, AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast gastric and

Db 1 CTCGCTAGTATGTTACATGCCCTAATGGTTCATTGCTAGACAGACTCTCTGGCTTG 60

CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63468 to AAB63721 and AAB63970
CC AAB63232 to AAB63468 to AAB63722 to AAB63970

Qy 278 AspGlyThrProProlLeuSerTrpHisMetArgCysIleIaIaGInGlyAlaIaAsn 297
Db 61 GATGTACTCCACCACTTCTTGGCATGAGATGCAAGATTGCTCAGGGTGCAGCTAAT 120

CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic

Dy 298 GlyIleasnPheluehiSGluasenh;sh;slleh;sargaspilleyseSeralaaahle 317
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 GGCTCATATTTTCTACATGAAATTCATCATATTTCATAGAGATATTAAGAAGCAATAATC 180

SQ Sequence 586 BP; 183 A; 103 C; 120 G; 173 T; 7 other;

[illegible]

Alignment Scores:
Pred No.:

338 LysPheAlaGlnThrValMetThrSerArgIleValGlnThrThrAlaTyrMetAlaPro 357

Score:	883.00	Matches:	1/6
Percent Similarity:	97.24%	Conservative:	0

Db 241 AAGTTGCCCAGACAGTCATGACTAGCAGAATTGTGGACAACAGCTTATATGCCACCA 300

Query Match:	37.01%	Indels:	2
	55		0

QY 358 GIWALAEWARGGLYGLUTILETHRPROLYSSERASPILETYRSERPHEGLYVALVALEU 377

ITS-10-001-254-16 (1-460) x 2AE22511 (1-586)

D5 301 GAGCTTTCGCTGGAGAAATACACCCCAATCTGATATTTACAGCTTTGGTGTGTTTA 360

258 LeuCysLeuVal]TyrVal]TyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeu 277

3/8 Leugidaleimigylleufoalavabwrgunhsaigylufooinlleulleu 39

Db 48 CTCGCTAGTATGTTACATGCCCTAATGGTTCATTGGCTAGACAGACTCTTGGCTTG 107

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818

QY 278 AspglyThrProProLeuSerTrpHisMetArgCysLysIleA1aG1nG1yAlaAlaAsn 297

Db 108 GATGTA CTCCACCACTTCTTGACATGAGATGCAAGATTGCTCAGGTCAGCTAAT 167

QY 417 smetAsnAspAlaAspSerThrSerValGluAlaMetTyr-SerValAlaSerGlnCys 436

298 G I Y I I E A B I P N E U C H I B G I W A B I M I B N I S I R E H I B A I G A B P I E L Y B S E I A I A S I I I E 31

Db 481 GATGATGATGCTGATTCACACTTCAGTTGAANCTATGTACCTCTGTTGNTAGTCAATGT 539

25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053

RESULT 13

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

ID AAF22511 standard; cDNA; 586 BP.

C: 338 T u d b o n] m o t h e r c o w e t] o y s] n] e t h e t h e r A] a t t y r M o t A] a B r o 357

AC AAF22511;
XX

288 AAGTTGCTCAGACATGACTAGCAGATTGTGGGAACAACAGCTTATATGCCACCA 347

DT 26-MAR-2001 (first entry)
XX

358 GluAlaLeuArg]vg]u]l]eThrProLysSerAsp]l]eTyrSerPheG]l]ValValLeu 377

DE human breast cancer associated antigen nucleotide sequence SEQ ID NO: 90.

CC	AAB95899	represent human amino acid sequences; and AAH13632 to AAH13632
CC	represent oligonucleotides, all of which are used in the exemplification	
CC	of the present invention.	
SQ	Sequence 577 BP; 163 A; 120 C; 105 G; 177 T; 12 other:	
 Alignment Scores:		
Pred. No.:	7,26e-78	Length: 577
Score:	854.00	Matches: 170
Percent Similarity:	93.48%	Conservative: 2
Best Local Similarity:	92.39%	Mismatches: 11
Query Match:	95.79%	Indels: 0
DG:	22	Gaps: 2
 US-10-001-254-16 (1-460) x AAH09161 (1-577)		
QY	219 ValAspIleThrTrcInduLuleuysGln-GlnPheAspGInguileuValMetAl	238
DB	550 GTTGCANCTNCTACATVAAAGACTTGAAACCAGCAGTTGTATCAAGAATN-AAAGTAATGGC	492
QY	238 alySvSGlInhiSGLuaSnLeuVaIGluDeuLeuglyPheSerSerAspGIyaSPApLe	256
DB	491 AAAGTCGAACAGAAAACCTTAGTAGAACAATTGGTTCTTCAGATGATGAGATGACT	432
QY	258 uCySLeuValTyrrValTyrmecProASnglySerleuLeuaspArgLeuSerCySLeuAs	278
DB	431 CTGCTTAGATATATGTTTACATGCCCTAACGGTTCAATGCTACAGACTCTTCTTGGA	372
QY	278 pGIyTHrProroleuSerTrPhieMeArGcybyrGelaeInglylaalaangl	298
DB	371 TGSTACTCCACCACTTCTTGCCACATAGAGAGCAAGATGCTCAGGGTGCAGCTAAATGG	312
QY	298 yllEanPheUhiSGLuaSnhiShiSlGhiSarGAspIllelySerAlaenillele	318
DB	311 CATCATTTTTCTCAACAGAAAATCATCATATTCATAGATATTTAAAGNCMAATATCTT	252
QY	318 uLeuAspGluAlaPheThrAlaLyIsleSerAspPheGlyLeuAlaaArgAlaSerGluLy	338
DB	251 ACTGATGGAAGCTTTTANTGCTTAAAFATATCGNCTTGGCTTGACGGGCTTWVGAA	192
QY	338 sPheAlaGInThrValMetThSerArgllevalGlyThrAlaTyrmelAProGl	358
DB	191 GTTTCGCCACAGATCATACATAGAGAAATGTGGGAACAACCTTATATGGCACAGA	132
QY	358 uAlaLeuAsxgylGluileThrProLySerAspIleTySerPheGlyValValleuLe	378
DB	131 AGCTTTCGCTGGAGAAATACACCCAAATTCGATATTTACAGCTTGTGTGTTTACT	72
QY	378 uGUllelelThrNglyLeuProAlaValaSPgluhisArgsluProGlnleuLeuAs	398
DB	71 AGAAATATATACGTGACTTCACGCTGTGATGAAACCGGTGAACCTGATTTATTCAGA	12
 RESULT 15		
ID	AAV63193 standard; cDNA; 483 BP.	
XX	AAV63193;	
XX	25-MAR-2003 (updated)	
DT	13-JAN-1999 (first entry)	
XX	cDNA from clone eh61_1 which encodes a secreted protein.	
XX	Secreted protein; immune stimulating; suppressing;	
KM	haematopoiesis regulating activity; tissue growth activity; activin;	
KM	inhibin activity; chemocactive; chemokinetic activity; haemostatic;	
KM	thrombolytic activity; anti-inflammatory activity; cadherin;	
XX	tumour invasion suppressor activity; tumour inhibition activity; ds.	

OS Homo sapiens.
 XX Key Location/Qualifiers
 FH CDS 85..450
 FT /+tag= a
 XX
 XX MO9844113-A1.
 XX
 XX 08-OCT-1998.
 PD
 XX
 PF 27-MAR-1998; 98MO-US06176.
 XX
 XX 28-MAR-1997; 97US-0823330.
 PR
 XX 25-MAR-1998; 98US-0047661.
 XX
 XX (GEMV) GENETICS INST INC.
 PA
 XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
 PI Racie LA, Spaulding V, Treacy M;
 PI
 XX WPI; 1998-542703/46.
 DR P-PSDB; AAW80409.
 XX
 XX New isolated polynucleotide(s) and secreted proteins - are obtained
 PT from human cDNA libraries prepared from adult testes, foetal brain,
 PT adult brain, adult blood and placenta
 XX
 XX
 PS Claim 25; Page 82; 124pp; English.
 XX
 CC The present sequence encodes a secreted protein. The nucleic acid
 CC sequence is isolated from a human adult blood cDNA library using
 CC probe AAV63204. The polypeptide may have biological activities such as
 CC e.g. nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, tumour
 CC inhibition activity or other activities.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 XX
 SQ Sequence 483 BP; 156 A; 84 C; 101 G; 142 T; 0 other;
 Alignment Scores:
 Pred. No.: 2.62e-70 Length: 483
 Score: 779.00 Matches: 159
 Percent Similarity: 98.76% Conservative: 0
 Best Local Similarity: 98.76% Mismatches: 1
 Query Match: 32.65% Indels: 2
 DB: 19 Gaps: 0
 US-10-001-254-16 (1-460) x AAV63193 (1-483)
 QY 270 LeuLeuAspArgLeuSerCysLeuAspGlyThrProProLeuSerThrPheMetArgCys 289
 DB 3 TTGCTAGACAGACTCTCTTCTTGATGATGATCTCCACCACTTCTTGACATGAGATGC 62
 QY 290 LysIleAlaGlnGlyAlaIleAsnGlyIleAsnPheLeuHisGlnAsnHisIleHis 309
 DB 63 AACATTGCTCAGGCTGAGCTTAAGGATCAATTTCTTACATGAATATCATCATATTCT 122
 QY 310 ArgAspIleLeuSerAlaAsnIleLeuLeuAspGluAlaPheThrAlaLysIleSerAsp 329
 DB 123 AGAGATATTTAAAGTCAATATCTTACTGATGAAGCTTTTACTGTAATATCTGAC 182
 QY 330 PheGlyLeuAlaArgAlaSerGlyLys-PheAlaGlnThrValMetThrSerArgIleVal 349
 DB 183 TTTGGCTTGCACGGGCTTCTGAGAGTTTGGCCACAGACATGATGATGAGCAATTGT 242
 QY 349 IGIYThrThrAlaTyMetAlaProGluAlaLeuArgGlyGluIleThrProLysSerAs 369
 DB 243 GGGAAACAGCGTTATATGACCAAGAGCTTGGCGTGAAGAAATTAACACCAAAATCTGA 302
 QY 369 PileTySerPheGlyValValLeuLeuGluIleIleThrGlyLeuProAlaValAspG1 389

DB 303 TATTTACAGCTTTGTGTGTGTTTACTAGAAATTAATTAAGTCCAGCTGTGATGA 362
 QY 389 UH1ARGGluProGlnLeuLeuAspIleLysGluGluIleGluAspGluGlySerTh 409
 DB 363 ACACCGTAACCTTCAGTTATTTGCTAGATATTTAAAGAAATTCAGATGAAGAAAGAC 422
 QY 409 rIleGluAspTyrlleAspLysLysMetAsnAspAlaAspSerThrSerValGluAlaMe 429
 DB 423 ATT-GAAGATTATATGATTAATAAAGATGAATGATGATTCACCTTCAGTTGAAGCTAT 481
 QY 429 t 429
 DB 482 G 482

Search completed: January 18, 2004, 01:32:15
 Job time : 567.702 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 18, 2004, 01:04:18 ; Search time 7784.21 Seconds
(without alignments)
2417.512 Million cell updates/sec

Title: US-10-001-254-16
Perfect score: 2386
Sequence: 1 MNKPTPTSTYVRCILNVGLIR.....KNRPDIKKVQQLQEMTAS 460

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO/US10001254/runat.16012004.152423.19714/app_query.fasta.1.1109
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ext -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10001254@cgn.1.1.5671@runat.16012004.152423.19714 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:
1: gb_ba:*
2: gb_hvg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_secs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_secs:*
28: em_un:*

29: em_vl:*
30: em_hvg_hum:*
31: em_hvg_inv:*
32: em_hvg_other:*
33: em_hvg_mus:*
34: em_hvg_pln:*
35: em_hvg_rod:*
36: em_hvg_mam:*
37: em_hvg_vrt:*
38: em_sy:*
39: em_hvgo_hum:*
40: em_hvgo_mus:*
41: em_hvgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2386	100.0	1383	6 AX431306	AX431306 Sequence
2	2386	100.0	1629	9 BC013316	BC013316 Homo sapi
3	2386	100.0	2820	9 AK000528	AK000528 Homo sapi
4	2382	99.8	1383	6 AX196260	AX196260 Sequence
5	2382	99.8	1383	9 AF45802	AF45802 Homo sapi
6	2362	99.0	2817	6 AX431318	AX431318 Sequence
7	2362	99.0	2817	6 AF155118	AF155118 Homo sapi
8	2017	84.5	1542	6 AX196262	AX196262 Sequence
9	2017	84.5	1542	10 AF45803	AF45803 Mus muscu
10	2016	84.5	2431	10 BC051676	BC051676 Mus muscu
11	1787.5	74.9	2213	6 BD155790	BD155790 Primer fo
12	1787.5	74.9	2213	6 AK027301	AK027301 Homo sapi
13	1169	49.0	833	6 AR223870	AR223870 Sequence
14	981	41.1	1719	5 BC045381	BC045381 Dario rer
15	883	37.0	539	6 AX053431	AX053431 Sequence
16	883	37.0	586	6 AX053324	AX053324 Sequence
17	854	35.8	577	6 BD151153	BD151153 Primer fo
18	691	29.0	501	6 AX121132	AX121132 Sequence
19	666	27.9	2069	3 AK116443	AK116443 Ciona int
20	590	24.7	3303	9 AY186092	AY186092 Homo sapi
21	590	24.7	118572	9 AC093012	AC093012 Homo sapi
22	590	24.7	165868	2 AC021719	AC021719 Homo sapi
23	590	24.7	309582	2 AC025567	AC025567 Homo sapi
24	574	24.1	1287	6 AX653723	AX653723 Sequence
25	528.5	22.2	66237	8 AB016892	AB016892 Arabidops
26	525.5	22.0	2643	8 BT000634	BT000634 Arabidops
27	525.5	22.0	2821	8 AY091785	AY091785 Arabidops
28	517	21.7	294	6 AX431296	AX431296 Sequence
29	515.5	21.6	3590	6 AR270530	AR270530 Sequence
30	515.5	21.6	3590	6 AX330494	AX330494 Sequence
31	515.5	21.6	3590	6 AX409575	AX409575 Sequence
32	515.5	21.6	3590	6 AY59730	AY59730 Sequence
33	515.5	21.6	3590	9 HUM11R	159730 Sequence 1
34	512.5	21.5	24696	2 AC123081	L76191 Homo sapien
35	512.5	21.5	252400	2 AC094859	AC123081 Rattus no
36	508.5	21.2	1899	6 AX506104	AC094859 Rattus no
37	506	21.2	1930	8 AY150495	AX506104 Sequence
38	506	21.2	2585	8 AY056294	AY150495 Arabidops
39	506	21.2	2815	8 AY224496	AY056294 Arabidops
40	505.5	21.2	1281	6 AX412672	AY224496 Oryza sat
41	505.5	21.2	1281	8 AY125559	AX412672 Sequence
42	505.5	21.2	1539	8 D88207	AY125559 Arabidops
43	505.5	21.2	1539	8 D88207	D88207 Arabidops
44	505.5	21.2	1588	8 AF428432	AF428432 Arabidops
45	505	21.2	3051	8 BT002512	AF428432 Arabidops

RESULT 1

ALIGNMENTS

AX431306
 LOCUS AX431306 1383 bp DNA linear PAT 28-JUN-2002
 DEFINITION Sequence 15 from Patent WO240680.
 ACCESSION AX431306
 VERSION AX431306.1 GI:21656175
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Pawlowski, K., Fiorentino, L., Godzik, A., Lee, S. H., Reed, J. C.,
 TITLE Novel death domain proteins
 JOURNAL Patent: WO 0240680-A 15 23-MAY-2002;
 BURNHAM INST (US)
 FEATURES
 SOURCE location/Qualifiers
 1..1383
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 1..1383
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD37279.1"
 /db_xref="GI:21656176"
 /translation="MNKPIPTSTYVRCINVLIRKLSDFIDPEGMKKLAVAIKKPSG
 DDRYNPHIRRFALLLOTGKSPTELLFDWCTNCTGDIADILLIONFPAPSLILP
 DAVEKNTANLPKRAITYQKQKMPFCOKDRILMTPVONBOSVMPSSSPENSLV
 SDTRFHSFSEYELKNVTNNFDERPI SVGNKMGSGFVVKVGVNMTVAVKLAA
 VDIITBELKQDFDEIRKMAKCOHENLVELFGSDGDDCLVYVMPNGSLDLRLC
 LDGTPPLSMHRCKIAOGANGINFLHNHHRIDISANILDEAFKLSIDFGLAR
 ASEKFAQVMTSRIVGTAYMAVPAIRGEITPKSDIYSGVVLLEITGLPAVDEHE
 POLLLDKERIEDEKTEIDYIDKKMDADSTSEVAVSVASQCLHEKKNRPIKKV
 QQLQLEMTAS"
 BASE COUNT 463 a 243 c 283 g 394 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4.99e-198 Length: 1383
 Score: 2386.00 Matches: 460
 Percent Similarity: 100.00% Conservatave: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-001-254-16 (1-460) x AX431306 (1-1383)
 QY 1 MetAsnLysPProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
 Db 1 ATGAACAAACCCATTAACCATCAATATGCGCTCCCTCAATGTTGACTAATTAGG 60
 QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPlyValLeuAlaValAlaIleLys 40
 Db 61 AAGGTGACAGATTTTATGATCCCTCAGAGAGATGGAAGATTAGCGTAGCTAATTAA 120
 QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgArgPheGlnAlaLeuLeu 60
 Db 121 AAACCATCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 61 GlnThrGlyLysSerProThrSerGlnLeuLeuPheAspTyrPlyValThrAsnCysThr 80
 Db 181 CAACCTGGAAGAAAGTCCCACTTCTGAATTAAGTCTGTTGACTGGGACACCAAAATTGACA 240
 QY 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPhePheAlaProAlaSerLeuLeu 100
 Db 241 GTTGATGATCTTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGTG 300
 QY 101 LeuProAspAlaValAlaProLysThrAlaAsnThrLeuProSerLysGlnAlaIleThrVal 120
 Db 301 CTCCCAAGATGCTGTCTCCAAAGATGCTAATCACTTCTTAAAGAAAGCTAATAACAGATT 360
 QY 121 GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140

Db 361 CAGCAAAAAGATGCTCTTCTGTGACAAAGACAGGACATTGATGACACCTGTGACGAAT 420
 QY 141 LeuGlnInserLysMetProProAspSerSerSerProGlnLysAsnLysLeuGlnVal 160
 Db 421 CTGGAACAAAGCTATATGACCACTGACTCTCAAGTCCAAATAATTAAGATTAGAGATT 480
 QY 161 SerAspThrArgPheHisSerPheSerPheTyrGlnLeuLysAsnValThrAsnAspPhe 180
 Db 481 AGGATATCAAGTTTTCACAGTTTTCATTTTATGAATGAAGATGCACAAAATTAATT 540
 QY 181 AspGlnArgProIleSerValGlyValAsnLysMetGlyGlnGlyValPheGlyValVal 200
 Db 541 GATGAACGACCACTTCTCTGTGTGTGAATTAATGGAAGAGAGAGATTTGAGATTGTA 600
 QY 201 TyrLeuGlyTyrValAsnAsnThrThrValAlaValLysLysLeuAlaIleMetValAsp 220
 Db 601 TATTAAGGCTACGTAATTAACAACTGTGCGAGTGAAGAGCTTGACGCAATGTTGAC 660
 QY 221 IleThrThrGlnGlnLeuLysGlnGlnPheAspGlnGlnLysValMetAlaLysCys 240
 Db 661 ATTACTACTGAAGACTGAAGACGACGTTGATCAAGAAATTAAGTAATGGCAAACTGT 720
 QY 241 GlnHisGlnLysLeuValGlnLeuLeuGlyPheSerSerAspGlyAspAspLeuCysLeu 260
 Db 721 CAACATGAAACCTTAGTAGAATCACTTGTCTCAAGTATGAGATGAGATGACCTTGCTTA 780
 QY 261 ValTyrValTyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThr 280
 Db 781 GTATATGTTTACATGCTTAATGTTGTTCACTTGACAGACTCTTCTTGATGATGTAAT 840
 QY 281 ProProLeuSerTrpHisMetArgCysLysIleAlaGlnGlyAlaAlaAsnGlyIleAsn 300
 Db 841 CCACCACTTCTTGCGACATGAGATGACATGATGCTCAGGTTGAGCTAATGGATCAAT 900
 QY 301 PheLeuHisGlnLysAsnHisIleHisIleArgAspIleLysSerAlaAsnIleLeuLeuAsp 320
 Db 901 TTCTCATGATAAATCATCATATTCATATGAGATATTAAGATGCAATATCTTACTGAT 960
 QY 321 GlnAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGlnLysPheAla 340
 Db 961 GAACCTTTTACTGCTAAATATTCATGATTTGGCTTGACGGGCTTGTGAAGATTGGCC 1020
 QY 341 GlnThrValMetThrSerArgIleValGlyThrThrAlaTyrMetAlaProGlnAlaLeu 360
 Db 1021 CAGACATCATGACTACAGAAATTGTGGGAACAACAGCTTATATGGACCAAGAACTTGG 1080
 QY 361 ArgGlyGlnIleThrProLysSerAspIleTyrSerPheGlyValValLeuLeuGlnIle 380
 Db 1081 CGTGAAGAAATTAACCCCAATCATGATTTTACAGCTTGTGTGTGTTTACTAGAAATA 1140
 QY 381 IleThrGlyLeuProAlaValAspGlnHisArgGlnProGlnLeuLeuLeuAspIleLys 400
 Db 1141 ATTAAGTACTCTCCAGCTGTGATGAAACACCGGAACTCTGATTTGCTAGATATTAA 1200
 QY 401 GlnGlnIleGlnLysGlnGlnLysThrIleGlnLysPyrIleAspLysLysMetAsnAsp 420
 Db 1201 GAAGAAATTTGAAGTGAAGAAAGAAAGACAAATTGAAGTATTATGATTAAGAAAGTGAATGAT 1260
 QY 421 AlaAspSerThrSerValGlnAlaMetTyrSerValAlaSerGlnCysLeuHisGlnLys 440
 Db 1261 GCTGATTCACCTTCACTGATGAAGCTATGATCTGTTGCTGATCAATGTCTGCAAGAAAG 1320
 QY 441 LysAsnLysArgProAspIleLysLysValGlnGlnLeuLeuGlnGlnMetThrAlaSer 460
 Db 1321 AAAAATTAAGAACCAAGACATTAAAGAGTTCAACGCTGTGCAAGAGATGACGCTTCT 1380
 RESULT 2
 LOCUS BC013316 1629 bp mRNA linear PRI 04-SEP-2001
 DEFINITION Homo sapiens, clone MGC:13330 IMAGE:4287014, mRNA, complete cds.
 ACCESSION BC013316
 VERSION BC013316.1 GI:15426431

KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 1629)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONTech Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 Contact: amaden@systemsbio.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Ketterman and Anuradha Madan
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
 Series: IRAL Plate: 19 Row: n Column: 24
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7705840.
FEATURES
 source
 1..1629
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:13330 IMAGE:4287014"
 /tissue_type="Brain, primitive neuroectodermal"
 /clone_lib="NIH MGC_56"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"
 71..1453
 /codon_start=1
 /product="unknown (protein for MGC:13330)"
 /protein_id="AAH13316.1"
 /db_xref="gi:15426432"
 /translation="MNKDIPTSPVRLNGLIRKLSDFIDPOEGMKKLAIVAKKPSGDDRYNOFIRREPLAQKSPSTSELDPMKTNCTVDLIDLIQNEFPAPSLILPDAVEKTNLTLSKALITVQOKOMPDCDRLMTLVQLDSITPDPSSSPENKSLVSDTRPHSFSEFLKNVTNPFERPLISVGNKMGEGFVVKGVNNTTVAVKLAAMVDITTEELKQDFDEIKYMAKCOHENLVELFGSDGDLCLVYVMPNGSLDLRLSLDGPPLSMHMRCKIAOGANGINLEHNNHHRDIKSAANTLDEAFPAKISDGLARASEKPAQVTMTSRIVGTAAVMAPALREIIPKSDISFGVILEITGLPANDVHREPOLIIDKEIEDEKTEIDYIDKMANDSTVEAMTSVASQCLHEKKRPPDKVQQLQEMTAS"
CDS
 BASE COUNT 542 a 291 c 328 g 468 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 6.18e-198 Length: 1629
 Score: 2386.00 Matches: 460
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 US-10-001-254-16 (1-460) x BC013316 (1-1629)
 QY 1 MetAsnLysProIleThrProSerThrTyValAlaGlyCysLeuAsnValGlyLeuIleArg 20
 Db 71 ATGAACAAACCATTAACACATCAATATGCGCTCCCTCAAGTGGACTAATTAGG 130
 QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPlyLysLeuAlaValAlaIleLys 40
 Db 131 AAGGTGAGATTTTATGATCTCCAGAGAGATGGAAGAGTTAGTGTAGCTATTAAA 190

QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgArgPheGluAlaLeuLeu 60
 Db 191 AAACATCGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 250
 QY 61 GlnThrGlyLysSerProThrSerGlyLeuLeuPheAspTyrGlyThrAsnCysThr 80
 Db 251 CAAACTGGAAAAAGCCCATCTTGAATTAATGCTGGTGGAGGAGCACAATGGACACA 310
 QY 81 ValGlyAspLeuValAspLeuLeuIleGlnGlnGlnPhePheAlaProAlaSerLeuLeu 100
 Db 311 GTGGTGAATCTTGGAGATCTTTTATCCAAATGAATTTTGTCTCCGCAAGCTTTTGG 370
 QY 101 LeuProAspAlaValProLysThrAlaAsnThrLeuProSerLysGluAlaIleThrVal 120
 Db 371 CTCACAGATGCTGTTCCCAAAACGCTTAATCACTACTCTTAAAGAGCTAATACACTT 430
 QY 121 GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140
 Db 431 CAGCAAAAACAGATGCCCTTCTGTGACAAAGACAGACATGATGACACCTGTGAGAA 490
 QY 141 LeuGlnGlnSerTyrMetProProAspSerSerSerProGluAsnLysSerLeuGluVal 160
 Db 491 CTGAAACAAAGCTATATGACACTGACTCTCAAGTCCAGAAATTAAGTTAGAAAGTT 550
 QY 161 SerAspThrArgPheHisSerPheSerPheTyrGluLeuLysAsnValThrAsnAspPhe 180
 Db 551 AGTATACACCTTTTACACAGTTTTCATTTATATGATTAAGAAATGTCACAAATACCTT 610
 QY 181 AspGluArgProIleSerValGlyLysAsnLysMetGlyGlnGlyPheGlyValVal 200
 Db 611 GATGAACGACCATTTCTGTTGTGTGTATTAATGGAGAGGAGAGATTTGAGTTGTA 670
 QY 201 TyrLysGlyTyrValAsnAsnThrThrValAlaValLysLysLeuAlaIleMetValAsp 220
 Db 671 TATAAGGCTACGTAATAATACACAACTGTGCACTGAAGAGCTTGACAGCAATGTTGAC 730
 QY 221 IleThrThrGlnGluLeuLysGlnGlnPheAspGlnGluIleLysValMetAlaLysCys 240
 Db 731 ATTACTACTGAAGACTGAACAGACAGCTTATCATCAAGAAATTAAGTAATGGCAAGCT 790
 QY 241 GlnHisGluAsnLeuValGluLeuLeuGlyPheSerSerAspArgLysAspAspLeuCysLeu 260
 Db 791 CAACATGAACCACTAGTAGAAGCTACTGTTGTTCAAGATGAGAGATGACCTCGCTTA 850
 QY 261 ValTyrValTyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThr 280
 Db 851 GTATATGTTTACATGCTTAATGATGTTCAATGCTTACAGACACTCTTGGCTTGATGTA 910
 QY 281 ProProLeuSerTyrPheIleMetArgCysLysIleAlaGlnGlyValAlaAsnGlyIleAsn 300
 Db 911 CCACCACTTTTGTGGACATGATGAGAAAGTCTCTCAGGGTGAAGCTTAATGGCAATCAT 970
 QY 301 PheLeuHisGluAsnHisHisIleHisArgAspIleLysSerAlaAsnIleLeuLeuAsp 320
 Db 971 TTTCTACATGAATATCATCATATTCATATGATATTAATATGAATATTTACTGTGAT 1030
 QY 321 GluAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGlnLysPheAla 340
 Db 1031 GAACTTTTACTGCTAAATATCTGACTTGGCTTGGCCAGCGGCTTGAAGAGTTGGCC 1090
 QY 341 GlnThrValMetThrSerArgIleValGlyThrThrAlaTyrMetAlaProGluAlaLeu 360
 Db 1091 CAGACAGTCATGACTAGCAGATTTGTGGAAACAACGCTTATATGAGCACCAAGAGCTTG 1150
 QY 361 ArgGlyGluIleThrProLysSerAspIleTyrSerPheGlyValAlaLeuLeuGluIle 380
 Db 1151 CGTGAGAAATTAACCAACCAATCTGATATTATACAGCTTTGGTGTGTTTACTGAATA 1210
 QY 381 IleThrGlyLeuProAlaValAspGluHisArgGluProGlnLeuLeuAspIleLys 400
 Db 1211 ATAACTGACATTCACAGCTGTGTGATGAACACCGTGAACCTCATTTATGCTAGATATA 1270

FEATURES	source	Location/Qualifiers
LOCUS	AK000528	2630 bp mRNA linear
DEFINITION	AK000528.1	GI:7020683
VERSION	AK000528.1	GI:7020683
KEYWORDS		oligo capping; f1s (full insert sequence).
SOURCE		Homo sapiens (human)
ORGANISM		Homo sapiens
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS		Watanabe, S., Kumagai, A., Itakura, S., Yamazaki, M., Taahiro, H., Ota, T., Suzuki, Y., Oba, T., Oba, Y., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, T., Isogai, T. and Sugano, S.
TITLE		NEDO human cDNA sequencing project
JOURNAL		2 (bases 1 to 2820)
REFERENCES		Unpublished
AUTHORS		Sugano, S., Suzuki, Y., Ota, T., Oba, Y., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, T., Isogai, T. and Sugano, S.
TITLE		Shibahara, T., Tanaka, T. and Nakamura, Y.
JOURNAL		Submitted (15-FEB-2000)
REFERENCES		Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
AUTHORS		NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' - 3' end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
COMMENT		Location/Qualifiers
		1..2820
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone="KAT10395"
		/cell_line="KATO III"
		/cell_type="signet-ring cell carcinoma"
		/clone_lib="KAT"
		/note="Cloning vector pME18SFL3"
		36..1412
		/note="unnamed protein product"
		/codon_start=1
		/protein_id="BAA91232.1"
		/db_xref="GI:7020684"
		/translation="MNKRIPSTYVRCINVLIRKSPDIPDGKMKLAVAIKKPSGDDRNQPIRRPFAALLQTKSPSTSELDKGTCTGVDLVDLQNEFPAPASLLFSDVETANTLPSSKEATIVQOKPCPCDRLMTLPVQLQESYMPDSSSPENSLLEAVDFHSPFELKRVNTNPFERPIPCGDNMGEGFQVYKGVNNTVAVKLAAVDTITELKQOFDOIKYMAKCHENLVEILGFSDDGDLCLVYVYNGSLDRLSCLIDGPPSLSMHRCKIAOGANGINFLHNHHIHDIXSANILLDEAPAKISDGLARASERPAQVYNTSRIVGTAVNAPALREITPKSDIYSPGVILFETGLPAVDENHREPOLLLIDKEIIEDEKCTIEDYIDKKMDADSTSEAMTSVASQCLHEKRRNPDKVQQLLEMTAS"
BASE COUNT		940 a 534 c 579 g 767 t
ORIGIN		

Percent Similarity:	2386.00	Matches:	460
Best Local Similarity:	100.00%	Conservative:	0
Query Match:	100.00%	Mismatches:	0
DB:	9	Indels:	0
Gaps:			0
US-10-001-254-16 (1-460) x AK000528 (1-2820)			
QY	1 MetAanlySProlIeThrProSerThrValArgCysLeuAsnValIglIleuIleArg	20	
Db	30 ATGAACAAACCCATACACACATCAATATGCGCTGCTCAATATGTGACATAATTAGG	89	
QY	21 LysLeuSerASPhePheIleASPProGlnGluGlyTTPlysLysLeuAlaValIleLys	40	
Db	90 AAGCTGTCAAGATTATTAATGATCCTCAAGAGAGATGAGAGAGATTAGCTGATTAATA	149	
QY	41 LysProSerGIYASPAPArgTYRAsnGlnPheIleAlaArgAPheGluAlaLeuLeu	60	
Db	150 AAACCATCGTGATGATATGATATCATCATGTTTCCATATAGAGATTGAGCATTAATT	209	
QY	61 GlnThrGIYLYSerProThrsSerGIuLeuLeuPheAspTIPGlyThrThrsenCysThr	80	
Db	210 CAAATCGGAAAAAGTCCACATTCGAATTACGTTTGATCGGGGCCCAAAATTCACACA	269	
QY	81 ValGIYAPLeuValASPLeuLeuIleGlnAsnGluPhePheAlaProXaIleSerLeuLeu	100	
Db	270 GTTGATGATCTGTGGATCTTTTGATTCAAAATAGATTTTGTCTCTCGAGATCTTTTG	329	
QY	101 LeuProASPAlaValProLYSThrAlaAsnThrLeuProSerLYSGluAlaIleThrVal	120	
Db	330 CTCCAGATGCGTGTCCCAAACTGCTTAATACATCACTTCTTAAGAGACTATACAGTT	389	
QY	121 GlnGlnLYSGIleMetProPheCysAspLYSAPArgThrLeuMetThrProValGlnAsn	140	
Db	390 CAGCAAAAAACAGATCCCTTTCTGTACAAAGACAGGACATTCATGACACCTGTGACAGAT	449	
QY	141 LeuGlnGlnSerTYMetProProAspSerSerSerProGluAsnLYSerLeuGluVal	160	
Db	450 CTTGAAACAAAGCTAATATGACCATCACTCTCAAGTCCGAAATAATAAAGTTTAGAGTT	509	
QY	161 SerASPThrArgPheIleSerPheSerPheTYRGIuLeuLYSAsnValThrAsnAsnPh	180	
Db	510 AGTGATACCGTTTTCACAGTTTTCATTTATGATTAAGAATATGTCACAAATTAATCTTT	569	
QY	181 AspGluArgProIleSerValGIYGIYAsnLYSmetGIYGIuGIYpHeGIYValVal	200	
Db	570 GATGAACGACCATTTCTGTGTGTGTATTAATAATGGGAGAGGAGATTGGAGTTGTA	629	
QY	201 TYTYLYSGIYTYRValAsnAsnThrThrValAlaValLYSLeuAlaIleMetValASP	220	
Db	630 TATAAAGGCTACGTAAATTAACCAACTGTGGGAGTGAAGAACCTTGACCAATAGTTTGAC	689	
QY	221 IleThrTrnGluGluLeuLYSGlnGlnPheAspGlnGluIleLysValIleMetAlaLysCys	240	
Db	690 ATTACTACTGAAGAACTGAAACAGACGTTTGATCAAGAAATTAATAGTAATGGCAAAGTGT	749	
QY	241 GlnHISGluAsnLYSValGIuLeuLeuGIYpHeSerSerASPGLYASPAPLeuCysLeu	260	
Db	750 CAACATGAAAACTTAGTGAATCACTTGTTCTCAAGTATGAGATGACCTCTGCTTA	809	
QY	261 ValTYRValTYRMetProAsnGIYSerLeuLeuAspArgLeuSerCysLeuAspGIYThr	280	
Db	810 GTAATGTATTCATCTCTTAATGTTCATATGCTAGACAGACTCTCTTGCTTGATGTGATCT	869	
QY	281 ProProLeuSerTTPHISMetArgCysLYSLeuIleAGlnGIYAlaIleAsnGIYIleAsn	300	
Db	870 CCACCACTTCTTGGCACATGAGATGACAGATTCGACGGGTGCACATATGCAATCAAT	929	
QY	301 PheLeuHISGluAsnHISHisIleHisArgAspIleLYSAsnLYSAsnIleLeuLeuASP	320	
Db	930 TTTCTACATGAAAAATCATCATATTTCTATGAGATATTAATAAGTGCAAAATCTTACTAGGAT	989	
QY	321 GluAlaPheThrAlaLysIleSerASPheGIYLeuAlaArgAlaSerGIYLYSPheAla	340	

DB 990 GAAGCTTTTACTGCTAAATATCTGACTTGCGCTTGACAGGCTCTGAGAACTTTGCC 1049
QY 341 GlnThrValMetTherSerArgIleValGlyThrThrAlaTyrMetAlaProGluAlaLeu 360
DB 1050 CAGACGATGATGATGACGAAATGTTGGAAACAAGCTTATATGACACCAAGCTTTGG 1109
QY 361 ArgGlyGluIleTherProIysSerAspIleTyrSerPheGlyValValIleuLeuGluIle 380
DB 1110 CGTGGAGAAATTAACACCAATCTGATATTACAGCTTTGGTGTGTTTACTAGAAATA 1169
QY 381 IleThrGlyLeuProAlaValAspGluHisArgGlyProGlnIleuLeuAspIleIys 400
DB 1170 AATACCTGAGCTTCCAGCTGTGATGAAACCGGAACTCAGTATATGCTAGATATAA 1229
QY 401 GlnGluIleGluAspGluGluIysThrIleGluAspTyrIleAspIleIysMetAsnAsp 420
DB 1230 GAAGAAATTAAGATGAAGAAAGACAAATTAAGATTATTAATAAAGATGAATGAT 1289
QY 421 AlaAspSerThrSerValGluAlaMetTyrSerValAlaSerGlnCysLeuHisGluIys 440
DB 1290 GGTGATTCACCTTCAGTTGAGCTATGATCTCTGTTGCTAGTCAATGCTGCATGAAG 1349
QY 441 LysAsnLysArgProAspIleIysIysValGlnGlnLeuLeuGlnIleuMetThrAlaSer 460
DB 1350 AAAAATTAAGACACGACATTAAGAGTTCAACAGCTCTGCAAGATGACAGCTTCT 1409
RESULT 4
AX196260
LOCUS AX196260 1383 bp DNA linear PAT 28-AUG-2001
DEFINITION Sequence 2 from Patent WO0151641.
ACCESSION AX196260
VERSION AX196260.1 GI:15384642
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Wesche, H. and Li, S.
TITLE Irak-4: compositions and methods of use
JOURNAL Patent: WO 0151641-A 2 19-JUL-2001;
Tularik Inc. (US)
FEATURES
source location/Qualifiers
1. 1383
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="human IL-1 receptor-associated kinase 4 (IRAK-4)
cDNA"
1. 1383
/note="human IRAK-4"
/codon_start=1
/protein_id="CAC60090.1"
/db_xref="GI:15384643"
/translation="NMKPTTSTVYRCINLVGLIRKLSDFIDPOBGMKLAVALIKKPSG
DDRVNOFHRRPRLALQNGKSPSTSLFPMWTNCTAGDINLLIIONEPRAASILIP
DAVRKNTATLPSKEATITVOQKMPFCDKDRITLTYQNTESGYMPPDSSSPNKSLAV
SDTRHFSFYLELKVNTNNFDERPIASVGNKMGEGFGVYGVYNNNTVAARKLAAM
VDITTEELKQPDFDEIKVMACQHNLELVEFGSSDGDLCIVVYMPNGSLDLRLSC
LDGTPPLSMHMKCKIAQGANGLINLEHNNHIDRIKSNLILDLDAFTAKISDFGLAR
ASEKFAQVMTSTRIIVGTATVYAPBALRGITPKSDIYSGVVALLEITGLPAVDHRE
POLLDIKEIEDEDEKTEIDYIDKGNDAIDSTVSAMSVASQCHKEKKRNPDIKY
QQLDQEMTAS"

BASE COUNT 463 a 244 c 283 g 393 t
ORIGIN
Alignment Scores:
Pred. No.: 1, 11e-197
Score: 2382.00
Percent Similarity: 99.78%
Best Local Similarity: 99.78%
Query Match: 99.83%
Length: 1383
Matches: 459
Conservative: 0
Mismatch: 1
Indels: 0

DB: 6 Gaps: 0
US-10-001-254-16 (1-460) x AX196260 (1-1383)
QY 1 MetAsnLysProIleTherProSerThrTyrValAlaArgCysLeuAsnValGlyLeuIleArg 20
DB 1 ATGAACAAACCATTAACACCATCAATATGAGCCCTGCTCAATGTGGACTTAATTAGG 60
QY 21 LysLeuSerAspPheIleAspProGlnGluIleTyrPheValLeuAlaValIleIys 40
DB 61 AAGCTGTGATTTTATGATCTCAGAGAGATGAGAGAGATGAGCTGTAGCTATTTAA 120
QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgArgPheGluAlaLeuLeu 60
DB 121 AAACCATCTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 61 GlnThrGlyLysSerProThrSerGluLeuLeuPheAspTyrGlyThrThrAsnCysThr 80
DB 181 CAACTGGAAAGGCTCCCACTCTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
QY 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPheAlaProAlaSerLeuLeu 100
DB 241 GCTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGTGAT 300
QY 101 LeuProAspAlaValProIysThrAlaAsnThrLeuProSerLysGluAlaIleThrVal 120
DB 301 CTCCAGATGCTGTTCCCAAACTGCTAATACACTACCTTCAAGAGCTATTAACAGTT 360
QY 121 GlnGlnLysGlnMetProPheCysAspIysAspArgThrLeuMetThrProValGlnAsn 140
DB 361 CAGCAAAAACAGATCCTTCTGTGACAAAGACAGACAGACAGATGATACCTGTGAGAT 420
QY 141 LeuGlnGlnSerTyrMetProProAspSerSerProGlnAsnLysSerLeuGluVal 160
DB 421 CTGGAACAAAGCTAATATGACACTGACCTCAAGTCAAGAAATTAAGATTGAAGTT 480
QY 161 SerAspThrArgPheHisSerPheSerPheTyrGluLeuLysAsnValThrAsnAsnPe 180
DB 481 AGTGATACACGTTTTCACAGTTTTCATTTATGATGATGATGATGATGATGATGATGAT 540
QY 181 AspGluArgProIleSerValGlyIysAsnLysMetGlyGluGlyPheGlyValVal 200
DB 541 GATGAACAGACCATTTCTGTGTGTGTGATATAAAGGAGAGAGAGAGATTGAGATTGTA 600
QY 201 TyrLysGlyTyrValAsnAsnThrThrValAlaValLysLeuAlaAlaMetValAsp 220
DB 601 TATTAAGGCTACGTAAATTAACCAACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 221 IleThrThrGluGluLeuLysGlnGlnPheAspGlnGluIleValMetAlaIysCys 240
DB 661 ATTACTACTGAAGAACTGAACAGCACTTGTATCAAGAAATTAAGTAAATGCAAGTGT 720
QY 241 GlnHisGluAsnLeuValGluLeuLeuGlyPheSerSerAspGlyAspAspLeuCysLeu 260
DB 721 CAACATGAAGAACTTGTAGAACTACTGTTGTTCTCAAGATGAGAGAGAGAGAGAGAG 780
QY 261 ValTyrValTyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspIleThr 280
DB 781 GATATGTTTACATGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 281 PropProLysSerTyrPheMetArgCysLysIleAlaGlnGlyValAlaAlaAsnGlyIleAsn 300
DB 841 CCACCACTTCTTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 301 PheLeuHisGluAsnHisIleIleHisArgAspIleLysSerValAsnIleLeuLeuAsp 320
DB 901 TTTCTACATGAAGAAATCATCATATTCATAGAGATATTAAGGCAAAATATCTTACGAT 960
QY 321 GlnAlaPheThrAlaIleLysIleSerAspPheGlyLeuAlaArgAlaSerGlnLysPheAla 340
DB 961 GAAGCTTTTACTGCTAAATATATGACTTGTGCGCTTGACAGGCTCTGAGAAAGTTTCC 1020
QY 341 GlnThrValMetTherSerArgIleValGlyThrThrAlaTyrMetAlaProGluAlaLeu 360

```

Db      1021 CAGACAGTCATGACATGAGAAATGTGGGAAACACAGCTTATATGGCAACCAAGCTTGG 1080
Qy      361 ArgGlyIuIleThrProIleSerAspIleTySerPheGlyValValIleuGluIle 380
Db      1081 CGTGAGAAATTAACACCAATTCATATTAACACCTTGGTGCTTTACTAGAAATA 1140
Qy      381 IleThrGlyLeuProAlaValIleAspGluIleSarGluProGlnIleuLeuAspIleIys 400
Db      1141 ATAACCTGACCTTCCAGCTGTGATGATACACCTGAACTCTGATATTCCTGATATTA 1200
Qy      401 GluGluIleGluAspGluGluIleThyThrIleGluAspTyIleAspIleIysMetAsnAsp 420
Db      1201 GAAAGAAATTAAGATGAAGAAAGCAATTAAGATTATTTGATTAAGAAAGATGATGAT 1260
Qy      421 AlaAspSerThrSerValGluAlaMetTySerValAlaSerGlnCysLeuIleGluIys 440
Db      1261 GCTGATTCACCTTGCCTGGAAGCTATGATCTCTGTGTGATGATGATGCTGATGAA 1320
Qy      441 LysAsnIysAspProIleThrProIleIleIysValGlnIleuLeuGlnIleMetThrAlaSer 460
Db      1321 AAAAATAGAGACGACATTAAGAAAGGTTCAACAGCTGTGCAAGATGACAGCTTCT 1380

RESULT 5
AP445802
LOCUS      1383 bp mRNA linear PRI 20-APR-2002
DEFINITION Homo sapiens interleukin-1 receptor associated kinase 4 (IRAK4)
ACCESSION AF445802
VERSION    AF445802
KEYWORDS   AF445802.1 GI:20219009
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 1383)
AUTHORS    Li,S., Strelow,A., Fontana,E.J. and Wesche,H.
TITLE      IRAK-4: a novel member of the IRAK family with the properties of an
            IRAK-kinase
JOURNAL    PROC. NATL. ACADE. SCI. U.S.A. 99 (8), 5567-5572 (2002)
MEDLINE    21957277
PUBMED     11960013
REFERENCE  2 (bases 1 to 1383)
AUTHORS    Suzuki,N., Suzuki,S., Duncan,G.S., Millar,D.G., Wada,T.,
            Mitsuo,C., Takada,H., Wakeham,A., Irie,A., Li,S., Penninger,J.M.,
            Wesche,H., Ohashi,P.S., Mak,T.W. and Yeh,W.C.
TITLE      Severe impairment of interleukin-1 and Toll-like receptor
            signalling in mice lacking IRAK-4
JOURNAL    NATURE 416 (6882), 750-756 (2002)
MEDLINE    21959395
PUBMED     11923871
REFERENCE  3 (bases 1 to 1383)
AUTHORS    Li,S., Strelow,A., Fontana,E.J. and Wesche,H.
TITLE      Direct Submission
JOURNAL    Submitted (08-NOV-2001) Biology I, Tularek Inc., 2 Corporate Drive,
            South San Francisco, CA 94080, USA
FEATURES   Location/Qualifiers
            1..1383
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            1..1383
            /gene="IRAK4"
            1..1383
            /gene="IRAK4"
            /note="protein kinase"
            /codon_start=1
            /product="interleukin-1 receptor associated kinase 4"
            /protein_id="AA015772.1"
            /db_xref="GI:20219010"
            /translation="MNKPTSPSTYVRGLNGLIRKLSDFIDPQGWKKLVAIVKPPSG
            DDRNQHIRFEALLTGKSPSELLFDWETTKTCTAGLDVLDLQNEFRFASLLIP
            DAVPKTANTLPSKEALTIVQKMPCKDKDRTLTMTVONLQSTYMPDSSSPENKSLV

```

```

SDTRFHSFSFYELKNTVNFDERPISVGNKMBEGGFVVYKGYVNTTVAVKLAAM
VDITTEELKQFDEIKVMKCOHENVELIGSSDDDCLYVYVNPNSLDRISC
LDGTPPLSMWRCKIQGAANGINFLHNHHRDIKSANILDEAFPAKISDFGLR
ASEKFAQVWTSRIVGTGTVAPADEALRGEPISDYSFGVLEIITGIPAVDEHE
POLIIDKERIBEBEKTIEDYIDKKNVDASTSVLEANYVASQGLHKKKRPDIKKY
QQLQENTAS"
BASE COUNT  463 a 244 c 283 g 393 t
ORIGIN
Alignment Scores:
Pred. No.: 1,11e-197 Length: 1383
Score: 2382.00 Matches: 459
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.83% Indels: 0
DB: 9 Gaps: 0
US-10-001-254-16 (1-460) x AF445802 (1-1383)
Qy      1 MetAsnIysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
Db      1 ATGAACAAACCCATTAACACCATCAATATGTGGCTGCTCAATGTTGGACTAATTAGG 60
Qy      21 LysLeuSerAspPheIleAspProGlnGluGlyThrIleIysLeuAlaValAlaIleIys 40
Db      61 AAGCTGCAGATTTTATTTGATCCTCAAGAAAGATGAAGAAAGTTAGCTGATGATTAA 120
Qy      41 LysProSerGlyAspAspArgTyAsnGlnPheHisIleArgArgPheGluAlaLeuLeu 60
Db      121 AAACCATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Qy      61 GlnThrGlyIysSerProIleSerGlnIleuLeuPheAspIleThrThrAsnCysThr 80
Db      181 CAAACCTGAAAGAAAGTCCACCTTCGAAATTCCTGTTGATGCTGGGCAACCAATTCACA 240
Qy      81 ValGlyAspIleuValAspIleuIleGlnAsnGlnPheIleAlaProAlaSerLeuLeu 100
Db      241 GCTGTGATCTTGGTGATCTTTGATGATCAAAAGAAATTTTGTCTCGGAGCTTTTG 300
Qy      101 LeuProAspAlaValProIleThrAlaAsnThrIleuProSerIysGluAlaIleThrVal 120
Db      301 CTCCCATGCTGCTTTCCCAAACTGCTTAATACACTTCTTAAGAACTTAATACACTT 360
Qy      121 GlnGlnIysGlnMetProPheCysAspIysAspArgThrIleuMetThrProValGlnAsn 140
Db      361 CAGCAAAACAGATGCTTTCTGATGACAAAGACAGGACATTTGATGACACCTGGCAGAA 420
Qy      141 LeuGlnIleSerTyMetProProAspSerSerSerProGlnIleuIleIysSerLeuGluVal 160
Db      421 CTGAAACAAAGCTATATGACCACTGACTCTCAAGTCCAAAGTCCAAAGTCTTAAGAGTT 480
Qy      161 SerAspThrArgPheHisSerPheSerPheTyGlnIleuIysAsnValThrAsnAspPhe 180
Db      481 AGCATATCACCTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCA 540
Qy      181 AspGluArgProIleSerValGlyIysAsnIysMetGlyGluGlyIysPheGlyValVal 200
Db      541 GATGAACGACCACTTCTGTTGGTGATTAATAATGGGAGAGAGATTGGAGTTGTA 600
Qy      201 TyIleIysGlyTyValAsnAsnThrThrValAlaValIysIleuAlaIleMetValAsp 220
Db      601 TATTAAGGCTTCACTTAATTAACACACACTGTGGCAGTGAAGAAAGCTTGGCAGAA 660
Qy      221 IleThrThrGlnIleuIysGlnIlePheAspGlnGluIleIysValIleMetAlaIysCys 240
Db      661 ATTACTACTAAGAAACGAAACGAGCTTTGATCAAGAAATTAAGATTAAGGCAAAAGTGT 720
Qy      241 GlnHisGluAsnIleuValGluIleuGlyPheSerSerAspGlyAspAspLeuCysLeu 260
Db      721 CAACATGAAACCTTAAGTAAGACACTTGTGTTCTCAAGTATGAGATGACCTCTGCTTA 780
Qy      261 ValTyValIlyMetProAsnGlySerLeuIleuAspArgLeuSerCysLeuAspGlyThr 280

```


Db 781 GATATATGTTTACATGCTTAAGTTCATTGCTAGACAGACTCTCTTGTGATGAGTACT 840
 Qy 281 ProProLeuSerTrpHisMetArgCysLysIleAlaGlnGlyAlaIleAsnGlyIleAsn 300
 Db 841 CACACCACTTCTTGGACATGAGATGCAAGATTGCTCAGGGTCACTTAATGACATCAAT 900
 Qy 301 PheLeuHisGluAsnHisIleHisMetArgAspIleLeuSerAlaAsnIleLeuLeuAsp 320
 Db 901 TTTCTATGATAAATATCATATATTATAGAGATTAATAAGGCAAAATATCTTACAGGAT 960
 Qy 321 GluAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGlyLysPheAla 340
 Db 961 GAAGCTTTTACCTGCTAAATATCTGACTTGGCTTGACAGGCTCTTCTAGAAAGTTTGGC 1020
 Qy 341 GlnThrValMetThrSerArgIleValGlyThrThrAlaTyrMetAlaProGluAlaLeu 360
 Db 1021 CAGACGATCATGACTGCGAATTTGGGAAACACAGCTTAATATGACACAGAGCTTTG 1080
 Qy 361 ArgGlyGluIleThrProLysSerAspIleTyrSerPheGlyValValLeuLeuGluIle 380
 Db 1081 CCGTGAAGAATAACACCAATCTGATATTACAGCTTTGGTGGTTTATCTAGAAATA 1140
 Qy 381 IleThrGlyLeuProAlaValAspGluHisArgGluProGlnLeuLeuAspIleLys 400
 Db 1141 ATTAAGTGAAGCTTCCAGCTGTGATGAGACACCGTGAACCTCAGTTATTTCTAGATTTAA 1200
 Qy 401 GluGluIleGluAspGluGluLysThrIleGluAspTyrIleAspLysLysMetLeuAsp 420
 Db 1201 GAAAGAAATTGAAGAGAAAGAAAGACAAATTAAGATTATTAAGATAAAGATGAATGAT 1260
 Qy 421 AlaAspSerThrSerValGluAlaMetTyrSerValAlaSerGlnCysLeuHisGluLys 440
 Db 1261 GCTGATTCACCTTCAGTTGAGCTAATGTAATCTCTGTTGCTGATCAATGTCATGATAAAG 1320
 Qy 441 LysAsnLysArgProAspIleLysLysValGlnGlnLeuLeuGlnGluMetThrAlaSer 460
 Db 1321 AAAAATAAGAGACAGACATTAAGAGGTTCAACAGCTGCTCAAGAGATGACAGCTTCT 1380
 RESULT 6
 AX431318
 LOCUS AX431318 2817 bp DNA linear PAT 28-JUN-2002
 DEFINITION Sequence 27 from Patent WO0240680.
 ACCESSION AX431318
 VERSION AX431318.1 GI:21656187
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homimidae; Homo.
 REFERENCE
 AUTHORS 1
 TITLE Pawloweki, K., Fiorentino, L., Godzik, A., Lee, S.H., Reed, J.C.,
 Roth, W. and Stenmer-Liwen, F.
 JOURNAL Novel death domain proteins
 Patent: WO 0240680-A 27 23-MAY-2002;
 BURHAM INST (US)
 FEATURES
 source Location/Qualifiers
 1..2817
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 50..1432
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD37285.1"
 /db_xref="GI:21656188"
 /translation="MNKPIPTSTYVRCINVLIRKLSDFIDPQSGMKLAVALAIKPSG
 DDRVNFHRRREALLOTGKSPSELSLFDWGTNTCTVDLVDLILQNFPAASLLP
 DAVPKTANTLPSEKALITVOQKMPGCDORTLMPVNLBOISYMPDSSPENSLEY
 SDTRHSPFELKNTVNNFDERPI SVGNGMGEGRPVYVCLVYVYNNPTVAVKLAAM
 VDIITBELKQGDDEIKYMAKQHENLYVELLQFSSDGDCLVYVYNNPTVAVKLAAM
 LDGTPLSMHWKCKIAQGAANGINLHNHHIHRDIKSANILLDLVAFNAKISDFGLAC
 ASEKPAQVMTSRIVGTTAYMAPEALRGEITPKSDIYSGVVLLEIITGLPAVDEHRE
 FOLLADIKEIEDEKTEIDYIDKMDADSTSVAMYSGASQCHKKNSPDIKKY

BASE COUNT 912 a 547 c 586 g 772 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,546-195 Length: 2817
 Score: 2362.00 Matches: 456
 Percent Similarity: 99.13% Conservative: 0
 Best Local Similarity: 99.13% Mismatches: 4
 Query Match: 98.99% Indels: 0
 DB: 6 Gaps: 0
 US-10-001-254-16 (1-460) x AX431318 (1-2817)
 Qy 1 MetAsnLysProIleThrProSerThrTyrValArgCysLeuAsnValGlyLeuIleArg 20
 Db 50 ATGAACCAACCCATTAACACCAATCAATATGATGCGCTGCTCAATGTGGACTAATTTAGG 109
 Qy 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrLysLysLeuAlaValAlaIleLys 40
 Db 110 AAGCTGATGAGATTTTATGATCTTCAAGAGATGGAAGAAAGTTAGCTATTTAA 169
 Qy 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgArgPheGluAlaLeuLeu 60
 Db 170 AAACCATCTGGATGATGATAGATACATCAATCAGTTTCAATAGAGAGATTTGAAGCATTTACT 229
 Qy 61 GlnThrGlyLysSerProThrSerGlnLeuLeuPheAspTyrGlyThrThrAsnCysThr 80
 Db 230 CAATCTGAAAAAGCCCACTTCTGAATTAATGTTGCTGGGGCACCACAAATTTGACCA 289
 Qy 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPhePheAlaProAlaSerLeuLeu 100
 Db 290 GTTGATGATCTTGTGATCTTTTATGATCCAAATGAATTTTGTCTCTGCGAGCTTTTG 349
 Qy 101 LeuProAspAlaValProLysThrAlaAsnThrLeuProSerLysGluAlaIleThrVal 120
 Db 350 CTCCAGATGCTGTGCTCCAAACGCTTAATACATCACTTCTTAAGAGCTATACAGATT 409
 Qy 121 GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140
 Db 410 CAGCAAAAAAGAGTCCCTTTCTGTGACAAACAGAGCAATTAAGATGACACCTGTGAGAT 469
 Qy 141 LeuGlnGlnSerTyrMetProProAspSerSerProGluAsnLysSerLeuGlnVal 160
 Db 470 CTGGAACAAAGCTTATATGCCACCTGACTCTCAAGCTCCAGAAATTAAGATTGAGAGTT 529
 Qy 161 SerAspThrArgPheHisSerPheSerPheTyrGlnLeuLysAsnValThrAsnAspPhe 180
 Db 530 AGTATACACGTTTTCACAGTTTTCATTTATGATTAAGAGATGTCACAAATTAATCTT 589
 Qy 181 AspGluArgProIleSerValGlyLysAsnLysMetGlyGlnGlyGlyPheGlyValVal 200
 Db 590 GATGACACCAACCATTTCTGTGTGTGATTAATAATGGAGAGGAGGATTTGGAGTTGTA 649
 Qy 201 TyrLysGlyTyrValAsnAsnThrThrValAlaValLysLysLeuAlaIleMetValAsp 220
 Db 650 TATTAAGCTTAGTAATATACACAACTGTGACAGAAAGCTTGCACATGAGTTGAC 709
 Qy 221 IleThrThrGlnGluLeuLysGlnGlnPheAspGlnGluIleLysValMetAlaLysCys 240
 Db 710 ATTACTACTGAAGAACTGAACACAGAGTTTCAATTAAGAAATTAAGTAATGCAAAAGCT 769
 Qy 241 GlnHisGluLeuLeuValGluLeuLeuGlyPheSerSerAspGlyAspAspLeuCysLeu 260
 Db 770 CAACATGAAAACTTAAGTAAGACTACTGCTTCCAAAGTATGAGATGACCTGCTGCTTA 829
 Qy 261 ValTyrValTyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThr 280
 Db 830 GTATATGTTTACATGCTTAATGATGTTTCAATGCTTAACAGACTCTTGTGCTGAGTACT 889
 Qy 281 ProProLeuSerTrpHisMetArgCysLysIleAlaGlnGlyAlaIleAsnGlyIleAsn 300
 Db 890 CCACCACTTCTTGGACATGAGATGACAAAGATTCTCAGGGTGCAGCTAATGACATCAAT 949

QY 301 PheLeuHISGLUaenHISISIEHISargSPILeYSerAlaenILeLeuLeuASP 320
 DB 950 TTCTCATGAAATCATCATATTCATAGAGATATTAAGATGCAATATCTTCTGAT 1009
 QY 321 GUAaAetherAlaValIleSerASPheGlyLeuAlaAGAlaSerGluLeuPheAla 340
 DB 1010 GAACCTTTACTGCTAAATATCTGACTTGGCCCTTGCACGGCTTCTGAAAGCTTTGCC 1069
 QY 341 GlnThrValMetThrSerArgIleValGlyThrThrAlaTyMetAlaProGluAlaLeu 360
 DB 1070 CAGACAGTCACTAGCTAGCAAGATTTGTGGAAACACAGCTTATATGGACCAAGACTTTG 1129
 QY 361 ArgGlyGluIleThrProLYSserASPLeTySerPheGlyValIleLeuLeuGluIle 380
 DB 1130 CGTGGAAATTAACACCAATCTGATATTACAGCTTGTGTGTTTACTGAATA 1189
 QY 381 IleThrGlyLeuProAlaValASPGLuHISargGluProGlnLeuLeuASPILeLYS 400
 DB 1190 ATAACTGCACTCCAGCTGTGGATGAACCCGTGAACCTCGATTATGCTAGATATATAA 1249
 QY 401 GlnGluIleGluASPGLuGluLeuThrIleGluASPTrpIleASPLeuLYSLeuMetASP 420
 DB 1250 GAAGAAATTAAGATGAAGAAACACATTAAGATTTATTTGATTAAGATGAATGAT 1309
 QY 421 AlaASPserThrSerValGluAlaMetTySerValAlaSerGlnCySLeuHISGLuLYS 440
 DB 1310 GCTGATTCACACTGAGTGAAGTATGATCTGCTGCTGCTAGCAATGTCGAGTGAAGA 1369
 QY 441 LysASPLeuArgProASPILeLYSLeuValGlnGlnLeuLeuGlnGluMetThrAlaSer 460
 DB 1370 AAAAATAGAGCCAGACATTAAAGAGTTTCCACAGCTGTGCAAGAGATGACAGCTTCT 1429

RESULT 7
 AF155118 2817 bp mRNA linear PRI 05-JAN-2000
 LOCUS Homo sapiens putative protein kinase NY-REN-64 antigen mRNA,
 DEFINITION complete cds.
 ACCESSION AF155118
 VERSION AF155118.1 GI:5360130
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2817)
 Scanlan,M.J., Gordon,J.D., Williamson,B., Stockert,E., Bander,N.H.,
 Jongeneel,V., Gure,A.O., Jager,D., Jager,E., Knuth,A., Chen,Y.-T.
 and Old,L.J.
 Antigenes recognized by autologous antibody in patients with
 renal-cell carcinoma
 Int. J. Cancer 83 (4), 456-464 (1999)

TITLE Antigenes recognized by autologous antibody in patients with
 renal-cell carcinoma
 JOURNAL Int. J. Cancer 83 (4), 456-464 (1999)
 MEDLINE 99438124
 PUBMED 10508479
 2 (bases 1 to 2817)
 Scanlan,M.J., Gordon,J.D., Williamson,B., Stockert,E., Bander,N.H.,
 Jongeneel,V., Gure,A.O., Jager,D., Jager,E., Knuth,A., Chen,Y.-T.
 and Old,L.J.

TITLE Direct Submission
 JOURNAL Submitted (28-MAY-1999) Ludwig Institute, Sloan-Kettering
 Institute, 1275 York Ave, New York, NY 10021, USA
 FEATURES
 source
 1..2817
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="1973/10-4"
 /cell_type="renal cell carcinoma"
 50..1432
 /codon_start=1
 /product="putative protein kinase NY-REN-64 antigen"
 /protein_id="AADA2884.1"
 /db_xref="GI:5360131"

CDS
 /product="putative protein kinase NY-REN-64 antigen"
 /protein_id="AADA2884.1"
 /db_xref="GI:5360131"

/translation="MNKPTTSTYRCLNVGLIRKISDPIPOEGKKLAVALIKRSQ
 DDVYNGFHLIRREFALLQOTKSPSELIFPMGTNCVGLVDLLINEFAPASLLLP
 DAVPKFHLIRREFALLQOTKSPSELIFPMGTNCVGLVDLLINEFAPASLLLP
 SDVRFSFSELYELKATNTVOOKMPFCDKRTMTPTVOLEOSYMPDSSPEKKSILEV
 VDIITREISFOELKATNTVOOKMPFCDKRTMTPTVOLEOSYMPDSSPEKKSILEV
 LDGTTPPLSMNRKTIQGAANGINPFIHNNHHRDKSNTILDEAFYAKISPGCLAR
 ASERFQATVTSRIVGTIYVMAPEALRGELTPESDIYSROVLELEITGIPADVERHE
 PQLLDIKKEIIEDEETIEDIYIDKNNADSTISVEAMYSGASQCRHKKNKSPDIKKY
 HQLLOEMTAS"

BASE COUNT 912 a 547 c 586 g 772 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.54e-195 Length: 2817
 Score: 2362.00 Matches: 456
 Percent Similarity: 99.13% Conservative: 0
 Best local Similarity: 99.13% Mismatches: 4
 Query Match: 98.99% Indels: 0
 DB: 9 Gaps: 0

US-10-001-254-16 (1-460) x AF155118 (1-2817)

QY 1 MetASPLeuProIleThrProSerThrTrpValArgCysLeuAsnValIleLeuLeuArg 20
 DB 50 ATGAACAAACCCAAACACATCAACATATGTGGCTCTCAATGTTGACTAATTAAGG 109
 QY 21 LysLeuSerASPheIleASPProGlnGluGlyTrpLYSLeuAlaValIleLYS 40
 DB 110 AACGTGCATTTTATGATCTTCAAGAAAGATGGAAGAGTTAGCTGATTAATAA 169
 QY 41 LysProSerGlyASPAspArgGlyTranGlnPheHISLeArgArgPheGluAlaLeuLeu 60
 DB 170 AAACATCTGTGATGATATATCAATCAGTTTCACATTAAGAGATTGGAAGATTAATT 229
 QY 61 GlnThrGlyLYSserProThrSerGluLeuLeuPheASPTrpGlyThrTranCysThr 80
 DB 230 CAAACTGGAAGAAAGTCCCACTCTGATTAATCTGTTGACTGGGACACACAAATGGACA 289
 QY 81 ValGlyASPLeuValASPLeuLeuIleGlnAsnGluPhePheAlaProAlaSerLeuLeu 100
 DB 290 GTTGGTATCTTGTGATCTTTTGATCCAAATGAATTTTCTCTCGAGATCTTTTG 349
 QY 101 LeuProASPAlaValProLYSThrAlaAsnThrLeuProSerLYSGluAlaIleThrVal 120
 DB 350 CTCACGATCTGTCTCCAAACCTGTAATACCTACTCTTAAGAGCTAATAACAGTT 409
 QY 121 GlnGlnLeuMetProPheCysASPLeuArgThrLeuMetThrProValGlnAsn 140
 DB 410 CAGCAAAAACAGATGCTTCTGTGACAAAGACAGACATTAATGACACTGTGCAAGAT 469
 QY 141 LeuGlnIleSerTyMetProProASPserSerProGluAsnLYSserLeuGluVal 160
 DB 470 CTGGAACAAAGCTATATGCACTGACTCTCTCAAGTCCAAATAATAATTATGAAGTT 529
 QY 161 SerASPThrArgPheHISserPheSerPheTyGlnLeuLYSAsnValThrAsnAsnPhe 180
 DB 530 AGTATACAGCTTTTCACTTTTCAATTTATGAATTGAAGAATGACAAATAACTTT 589
 QY 181 ASPGluArgProIleSerValGlyLYSAsnLYSLeuGlyIleGlyPheGlyValVal 200
 DB 590 GATGAACGACCCATTTCTGTGTGTGATTAATAATGGAGAGAGATTTGGAGTTGTA 649
 QY 201 TyrLeuGlyTyTrValAsnAsnThrThrValAlaValLYSLeuAlaAlaMetValASP 220
 DB 650 TATTAAGGCTTACGTAATTAACACACCTGGCAGTGAAGAAGCTTGCAGCAATGCTTAC 709
 QY 221 IleThrThrGlnGluLeuLYSglnGlnPheASPglnGluIleLYSValMetAlaLYSCys 240
 DB 710 ATTACTACTGAAGAAGTGAAGACAGCTTTGATCAAGAAATTAATAAGTAAAGTGT 769
 QY 241 GlnHISGLUaenLeuValGluLeuLeuGlyPheSerSerASPLeuAspLeuCysLeu 260
 DB 770 CAACATGAAAACTTAGTAAGAACTTGTCTCAAGTATGATGAGATGACCTTGTCTTA 829

Oy		261	GATGTTTACATGGCTTAATGCTTCATTGCTGAACAGACTCTTGCTGGATGTACT	889
Oy		281	ProProlSerThrHisMetArgCysIlysLealaGlnGlyAlaAAsnGlyLeasn	300
Db		890	CCACCACCTTCTTGGCACATGAGATGCAAGATTGCTCAGGGTGCAGCTAAATGGCATCAT	949
Oy		301	PheLeuHISGLuaSerHIHisIleHisArgAspIleLysSerAlaAsnIleLeuleuAsp	320
Db		950	TTCCTACATGAAAACATCATTTTCATGAGATATTTAAAGTCAAAATACTTACTGCGAT	1009
Oy		321	GIuaIAPhetrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGluLysPheLea	340
Db		1010	GAGCTTTTACTGCTTAATAATATCTACCTTGCGCTTGACGGGGCTTCGAGAAGTTTCC	1066
Oy		341	GlnThrValMetThrSerArgIleValGlyThrThrAlaTyMetAlaProGluAlaLeu	360
Db		1070	CAGACAGTCATGACTAGCAGAAATGTGGGAACAACAGCTTAAATGCAACCGAACCTTGT	1122
Oy		361	ArgGlyGluIleThrProLysSerAspIleTyTyrSerPheGlyValIvalLeuLeuGluIle	380
Db		1130	CGTGAGAGAAATPAACCCAAATCTGATATTACACCTTGGTGCTTTTACTAGAAATA	1189
Oy		381	IleThrGlyLeuProAlaValAspGluHisArgGluProGlnLeuLeuLeuAspIleLys	400
Db		1190	ATACTGACATTCACGCTGCTGGATGAACAACCGTAGACCTCAAGTATTCCTGAATATTA	1249
Oy		401	GluGluIleGluAspGluGluLysThrIleGluAspTyrIleAspLysLysMetAsnAsp	420
Db		1250	GAGCAAATTTGAAGTGAAGAAAAAGCAATTTGAACATTAATTTGATTAATAAGATGAATAT	1309
Oy		421	AlaAspSerThrSerValGluAlaMetTyrSerValAlaSerGlnCysLeuHISgluLys	440
Db		1310	GCTGATTCACCTTCAGTGAAGCTATGTACTCTGGTGTAGCAATGTCCGATGTAAGAAAG	1369
Oy		441	LysAsnLysArgProAspIleLysLysValGlnGlnLeuLeuGlnGluMetThrAlaSer	460
Db		1370	AAAAATAAGACCAGACATTAAGAGAGTTCAACGAGCTGTGCANAGATGACAGCTTCT	1429
RESULT 8				
AX196262		1542 bp	DNA	linear PAT 28-AUG-2001
LOCUS				
DEFINITION		Sequence 4 from Patent WO0151641.		
ACCESSION		AX196262		
VERSION		AX196262.1	GI:15386464	
KEYWORDS				
SOURCE		Mus sp.		
ORGANISM		Mus sp.		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE				
AUTHORS		Wesche,H. and Li,S.		
TITLE		Irak-4 : compositions and methods of use		
JOURNAL		Patent: WO 0151641-A 4 19-JUL-2001;		
		Tularik Inc. (US)		
FEATURES				
source		location/Qualifiers		
		1..1542		
		/organism="Mus sp."		
		/mol_type="genomic DNA"		
		/db_xref="taxon:10095"		
		/note="murine IL-1 receptor-associated kinase 4 (IRAK-4) cDNA"		
		163..1542		
		/note="murine IRAK-4"		
		/codon_start=1		
		/protein_id="CAC6009.1"		
		/db_xref="gi:15386465"		
		/translation="MNKPILPSTYIRNLNVGLIRKLSDPTIDPGSKKKIAVAIKPKGS		
		DDRRNQHRIERFALLLOTGKSPRTCELLFDPMGTNTCYGDVLDVLQLELRPATLLIP		
		DAVEQTYSLEPRRAIVAQGHGCHQEDRTSVNPMPLDHSCEPDPSSSPDNRSVS		
		SDTRFHSPFSFELKSIITNNPEQPASAGNNMGGGCVGYKGVCNVTIVA VKLGAM		
CDS				

BASE COUNT	421	a	423	g	306	t
ORIGIN	Q0LL0E5A"					

Alignment Scores:	
Pred. No.:	6.82e-166
Score:	2017.00
Percent Similarity:	90.85%
Best local Similarity:	83.98%
Query Match:	84.53%
DB:	6
Length:	1542
Matches:	385
Conservative:	32
Mismatches:	42
Indels:	0
Gaps:	0

US-10-001-254-16 (1-460) x AX196262 (1-1542)

Qy	1	MeAsnAspProIleThrProSerThrTyrValArcCybLeuAsnValGlyLeuIleArg	20
Db	163	ATGACAAAGCCGCTTGACACCATCGACATACACGCAACTTAATGTGGGGATCTTAGG	222
Qy	21	LysLeuSerAspPheIleAspProGlnGluGlyTyrPylsValLeuAlaValIleLys	40
Db	223	AACTGTGGATTATTATTGATCTCTCAAGAAAGGGGTGGAAAGAAATTAGCAGTAGCTATCAA	282
Qy	41	LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgArgPheGluAlaLeuLeu	60
Db	283	AAACCGTCGGCCACACACAGATCAATCAAGTTCCATATAGAGGATTCGAAGCCCTTAATT	342
Qy	61	GlnThrGlyLysSerProThrSerGlnLeuLeuPheAspTyrPylsTyrThrAsnCysThr	80
Db	343	CACACCGGGAAGAGCCCACTGCTGAACTGCTTTTGAACGGGGCACACGAATCGCAC	402
Qy	81	ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPhePheAlaProAlaSerLeuLeu	100
Db	403	GTTGGGAGACCTGTGTGATCTACTACGATTCAGATTGACCTGTTTCCCGCCACACTCTCTG	462
Qy	101	LeuProAspAlaValProlysthrAlaAsnThrLeuProSerArgGluAlaIleThrVal	120
Db	463	CTCCCGGATGCGCTTCCCAAAACGTCGCAAAAGCTCTCTCTTAGAAGAAAGGGCAACAGTG	522
Qy	121	GlnGlnLysGlnMetProPheCysAspPylsAspArgThrLeuMetThrProValGlnAsn	140
Db	523	GCACAAACACACGGCGCTTGTACAGAAAAGACAGACATCCGTAATGCTATGCCGAAG	582
Qy	141	LeuGlnGlnSerTyrMetProAspSerSerSerProGluAsnLysSerLeuGluVal	160
Db	583	CTGAGAACACAGCTGCAGACCCGACCTCTCAAGCCACAGACAAACGAAGTGAAGTCC	642
Qy	161	SerAspThrArgPheHisSerPheSerPheTyrGlnLeuLysAsnValThrAsnAspHe	180
Db	643	AGGCAACACTCGTTCCACAGCTTCTGTTCCAAAGACTGAAGGCACTCAACAAACATTC	702
Qy	181	AspGluAspProIleSerValGlyGlyAsnLysMetGlyGluGlyPheGlyValVal	200
Db	703	GACGACCAACCCCGCTCTCCGCGTGGCAACCGGATGGAGAGGGGGGATTTGGAGTGGTG	762
Qy	201	TyrLysGlyTyrValAsnAsnThrThrValAlaValLysLysLeuAlaIleMetValAsp	220
Db	763	TACAAAGCGCTGTGTGAACAAACACATCGTGGCGGTAAAGAAAGCTCGACGAGTGTGAA	822
Qy	221	IleThrThrGlnGlnLeuLysGlnGlnIlePheAspGlnGluIleValMetAlaLysCys	240
Db	823	ATCAGTACGGAAGAACTAAAGCAACGTTTGATCAAGAAATTAAAGTAATGCGCAACGTGT	882
Qy	241	GlnHisGluAsnLeuValGlnLeuLeuGlyPheSerSerAspGlyAspAspLeuCysLeu	260
Db	883	CAGCAGAGAACTGTGTGAGCTGCTCGGCTTCTCCAGGACAGAGCAACACTGTGCTTA	942
Qy	261	ValTyrValTyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThr	280
Db	943	GGGTAAAGCTTAATGCCCAACGGATCTTGTCTGGACAGACTGTCTCCGCTGGATGGTAA	1002

Qy 281 ProProleusSerrTPHISmeCArgCyselysilealaglnlglalaaasnglytleasn 300
Db 1003 CCACCGCTTCTCTGGCACAAGGTCAGAGGTTGCTCAGGGGACAGCAAAATGGCATCAGG 1062
Qy 301 PheleuhsiglueasnhisilehiEargapPilleysSerAlaAenlileuleuasp 320
Db 1063 TTTCTGCATGAAATATCATCATCATTCATGAGATATTTAAATATGCAAAATATCTTACTAGAC 1122
Qy 321 GluAlaPheThrAlaValIleSerAspPheGlyleuAlaAsgAlaSerGluLysPheAla 340
Db 1123 AAAGACTTACTGCCAAATATCTGACTTGTGGCTTGACAGGGCTTCCGCAAGGCTACCG 1182
Qy 341 GluThrValMetThrSerArgIleValGlyThrThrAlaTyrmecAlaProGluAlaLeu 360
Db 1183 CAGACGGTCATGACACGCGCATCGTGGCACAACGGCTTACATGGCACCCGAAGCTTGG 1242
Qy 361 ArgGlyGluIleThrProLysSerAspIleTySerPheGlyValValLeuLeuGluIle 380
Db 1243 CGGGAGAAATTAACACCCAAATCTGACATCTACAGCTTCGGCGTGTCTGTGGAGCTG 1302
Qy 381 IleThrGlyLeuProAlaValAlaSpGluHisArgGluProGlnLeuLeuAspIleLys 400
Db 1303 ATAAACCGGCTGGCGCTGTGATGAATAACCGTGAACCTCAACTACTGCTGATATTAA 1362
Qy 401 GluGluIleGluAspGluGluLysThrIleGluAspTyrlleAspLysLysMeCAsnAsp 420
Db 1363 GAAGAGATTGAAGTGAAGAGAAGACGATTCAGAGATTACAGGATGAGAATATAGCAT 1422
Qy 421 AlaAspSerThrSerValGluAlaMetTySerValAlaSerGlnCysLeuHisGluLys 440
Db 1423 GCGGACCTCTGCTGCTGAGAGCATGTACTCTGCTGACAGCTGTCTCATGACAGAG 1482
Qy 441 LysAsnLysAspProAspIleLysLysValGlnGlnLeuLeuGlnGluMetThrAla 459
Db 1483 AAAACAGACGCGCCAGACATTCAGAAAGGTTCACAGCTGTCAAGAGATGTCTGCT 1539

RESULT 9
AF445803 1542 bp mRNA linear ROD 20-APR-2002
LOCUS
DEFINITION Mus musculus interleukin-1 receptor associated kinase 4 (Irak4)
ACCESSION AF445803
VERSION AF445803.1 GI:20219011
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Li, S., Strelow, A., Fontana, E. J., and Wesche, H.
TITLE IRAK-4: a novel member of the IRAK family with the properties of an
IRAK-kinase
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (8), 5567-5572 (2002)
MEDLINE 21957277
PUBMED 11960013
REFERENCE
AUTHORS Li, S., Strelow, A., Fontana, E. J., and Wesche, H.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-2001) Biology I, Tularik Inc., 2 Corporate Drive,
South San Francisco, CA 94080, USA
FEATURES
source
1..1542
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
1..1542
/gene="Irak4"
163..1542
/gene="Irak4"
/note="protein kinase"
/codon_start=1

/product="interleukin-1 receptor associated kinase 4"
/protein_id="AA015773.1"
/db_xref="GI:20219012"
/translation="MKKPLPSTYIRINLVNGLRKLSDFIDPOEGMKLVAIKPSG
DPRNOFHRRFEALLQTSKPTCELLFPMGCTNRCVGLVDLVLNVEIAPATLLP
DAVPOYKSLPPEALVAVOTGSPCEKORTSMPMPKLEHSECEPDSPPDRSVS
SDTRFHSFSEHLKSTINNEDEQPSAGARMEGGGVKKCVNNTVAVKKLGAM
VEISTEBELKQFPQEI KVMATCOHENVELGFSSDSNICLVAYAWPNSLDRLSC
LDGPTPLSMTRCKVAGTANIGIRFLPHENHIRDKSANILDKDFTAKISDGLAR
ASARIQVWTSRI VGTGTANMAPEALRGETTPKSDYSFGVVLLEITGLAAYDNE
POLLDIKERIEDEKTIEDYTDENKSDADPASVEAMYSASGCLHKKNRRIADY
QQLQEKSA"
BASE COUNT 421 a 392 c 423 g 306 t
ORIGIN
Alignment Scores:
Pred. No.: 6,82e-166 Length: 1542
Score: 2017.00 Matches: 385
Percent Similarity: 90.85% Conservative: 32
Best Local Similarity: 83.88% Mismatch: 42
Query Match: 84.53% Indels: 0
DB: 10 Gaps: 0
US-10-001-254-16 (1-460) x AF445803 (1-1542)
Qy 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
Db 163 ATGAACAGCGCTTGAACACATGACATGACATACATGACACTTATATGCGGAGATCCTTAGG 222
Qy 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrlleLysLysLeuAlaValIleLys 40
Db 223 AAGCTGCGGATTTATTTGATCTCTCAAGAGGCTGGAAGAAATATGACGATATCAAA 282
Qy 41 LysProSerGlyAspAspArgTyraGlnPheHisIleIleArgPheGluAlaLeuLeu 60
Db 283 AAGCGTCCGCGAGACGACATACATCACTTCATTAAGAGATTCGAAGCTTACTT 342
Qy 61 GluThrGlyLysSerProThrSerGluLeuLeuPheAspTyrlleThrAsnGlyThr 80
Db 343 CAGACCGGGAAGAGCCGACCATGTAAGCTGTTGACCTGGGGCACACAGAACTGCACA 402
Qy 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGluPheAlaProAlaSerLeuLeu 100
Db 403 GTTGGCAGCTTGTGGATCTTACTGCTCCAGATTGAGCTGTTTCCCGCCGACTCTCTG 462
Qy 101 LeuProAspAlaValProLysThrAlaAsnThrLeuProSerGlyAlaIleThrVal 120
Db 463 CTGCGGATGCGCTTCCCAACCGTCAAAAGCTGCTCTTGAAGAGCGGCAACGCTG 522
Qy 121 GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140
Db 523 GCACAAACACACGCGGCTTGTACGAAAGAGACAGACATCCGTAATGCTATGCCAAG 582
Qy 141 LeuGlnLysSerTyrmecProProAspSerSerSerProGluAsnLysSerLeuGluVal 160
Db 583 CTAGAAACACAGCTGCGAGCAGCCGAGCTCTCAAGCCCAACAGAACTGTAGAGTCC 642
Qy 161 SerAspThrArgPheHisSerPheSerPheTyGluLeuLysAsnValThrAsnAspPhe 180
Db 643 AGCAGACCTGGTCCACACCTTCTGTCATGAAGTGAAGACATCACAAACAACTTC 702
Qy 181 AspGluArgProIleSerValGlyLysAsnMetGlyGluGlyGlyPheGlyValVal 200
Db 703 GACGAGCAACCCGCGTGTCCGCGTGGCAACCGATGGAGAGAGGGGATTTGAGTGTG 762
Qy 201 TyrlleGlyTyValaAsnThrThrValAlaValLysLysLeuAlaValMetValAsp 220
Db 763 TACAAGGCTGTGTGAACAACATGCTGCGGTGAAGAGCTCGGAGCGATGTTGAA 822
Qy 221 IleThrThrGluGluLeuLysGlnGlnPheAspGlnGluIleLysValMetAlaLysCys 240
Db 823 ATCAGTACTGAAGAACTTAAGCAACAGTGTGATCAAGAAATTAAGTAAGCAACGTGT 882

```

Qy 241 GlnhsgluasnleuvalgluleuenglpheserSeraspolyAspaspheucysleu 260
Db 883 CAGCAGAGAACTGGTGGAGCTGCTCGCTTCTCCAGGAGACAGACCACTGGCTTA 942
Qy 261 VAllyrValyMeProasnolysleuleuAspaspheucysleuaspolythr 280
Db 943 GTGTATGCTTACATGATGCCCAACGGGTCTCTGTCGACAGACTGCTCGCATGTATCA 1002
Qy 281 ProPoleuSerTriphHisMetArgCyslysllealagllyAlaAlaasnlylleasn 300
Db 1003 CCAACGCTTCTCTGTCGACCAAGGTGCAAGTGTCTCAGAGGAGACCAATGACATCAGC 1062
Qy 301 PheleuhsigluasnhsisilleshisargaspilleySerSeraAlaasnilleuasp 320
Db 1063 TTTTCGATGAAATATCATCATCTAGATATTAAAGCAATATCTTACTAGAC 1122
Qy 321 GluaAlpheThAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 340
Db 1123 AAAAGCTTACTGCAAAATATCTGCTTGGGCTTGCACGGGCTTGGCAAGGCTAGCG 1182
Qy 341 GluThrValMetThSerArgIleValGlyThrThrAlaTyMetAlaProGluAla 360
Db 1183 CAGAGCGTATACACGCGGAATCGTGGCAACAGCTTACATGCGACCCGACGCTTGG 1242
Qy 361 ArgGlyGluIleThProlysSeraspilleyTySerPheGlyValValleuenglule 380
Db 1243 CGGGAGAAATTAACACCCAAATCTGACATCTACAGCTTGGGCTTGTGTGAGCTG 1302
Qy 381 lletHrGlyLeuProAlaValaspGluHisargGluProGlnleuLeuaspilley 400
Db 1303 ATAAACCGGGCTCGCGCTGTGATGATAAAACCTGAACCTCACTACTGCTGATTTAA 1362
Qy 401 GluGluIleGluAspGluGluIleThrIleGluAspTyrlleAspIleTyMetAsp 420
Db 1363 GAGAGATTTGAAGATGAAGAGAGAGATTTGAAGATTTACAGATGAAGATGAGCAT 1422
Qy 421 AlaAspSerThSerValGluAlaMetTySerValAlaSerGlnCysleuhsigluys 440
Db 1423 GCGGACCCCTCGCTCGGTGAAGCATCTACTCTGCTGCTAGCCAGTGTCTGATGAGA 1482
Qy 441 LybaAnlysaRProaspilleylysllyValGlnGlnleuLeuGlnGlnMetThAla 459
Db 1483 AAAAAGAGAGCGCCAGACATTTCAAAAGTTCAACAGCTGCTCAAGAGATGCTCT 1539

```

RESULT 10
BC051676

LOCUS BC051676 2431 bp mRNA linear ROD 10-JUN-2003
DEFINITION Mus musculus interleukin-1 receptor-associated kinase 4, transcript variant 1, mRNA (CDNA clone MGC:60994 IMAGE:30017484), complete cds.

ACCESSION BC051676
VERSION BC051676.1 GI:30354527
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerogomphi; Muridae; Murinae; Mus.
1 (bases 1 to 2431)

Strausberg,R.V., Feringold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,K., Shemen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heish,F., Ditchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Uscin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mallah,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,

Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalek,U., Smalins,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL MEDLINE
PUBMED 12477932
2 (bases 1 to 2431)
Strausberg,R.
Direct Submission
Submitted (30-APR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Minoru Ko
CDNA Library Preparation: Yulan Piao and Minoru Ko
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louissege, H., Kowis, C.R., Sneed, A.O., Martin, R.G., Muzny, D.M., Nantavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAX Plate: 111 Row: 1 Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23943897.

FEATURES
source
1..2431
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="MGC:60994 IMAGE:30017484"
/tissue_type="Embryo, whole, mouse, 7.5-dpc"
/clone_lib="NIA Mouse 7.5-dpc Whole Embryo CDNA Library (Long)"
/lab_host="DH10B"
/note="Vector: pSPORT1"
1..2431
/gene="Irak4"
/note="synonyms: NY-REN-64, 9330209D03R1k"
/db_xref="LOCUSID:266632"
/db_xref="MGI:2182474"
168..1547
/codon_start=1
/product="interleukin-1 receptor-associated kinase 4, isoform a"
/protein_id="AAH51676.1"
/db_xref="GI:30354528"
/db_xref="LOCUSID:266632"
/translation="MNKPLTPSTYIRLNVGILRLKLSDFIDPDEGKKLAIVAIPKPSGDDRNOHIRFEALLOTGSPCELEFDGNTCTGVDLVDLVOLELPAPVLLLPDAVQATYKSLPPRAATVQTHGPGCKDSTSWMPKLTSHSCPPDSSPDNSVSVSVDTRFHSFHEHLKSTINNDSDOPASAGNRMGCGGCVYKGVNNTTYAVKLGAMVISTEELKQDFDEILVMAVTCQENHVELLGFSSDSDNCLVYAVMPSGLDLRLSCLDGPPISMTIRCKVAGGTANGRIFLHNHHIHRDIDISANILDRPTAKISDGLARASALAVTWTSTRIVGTAAVMAPEALGEITPKSDIYSFVAVLELITGLAAVDENEPOLLIDKEERIEDEKTIETYDEKMSDADPASVEAMYSAASQCLHEKORRDIAYVQQLQEMSA"

BASE COUNT 656 a 580 c 618 g 577 t
ORIGIN
Alignment Scores:
Pred. No.: 1.51e-165 Length: 2431

Score: 2016.00 Matches: 385
 Percent Similarity: 90.63% Conservative: 31
 Best Local Similarity: 83.88% Mismatches: 43
 Query Match: 84.49% Indels: 0
 DB: 10 Gaps: 0

US-10-001-254-16 (1-460) x BC051676 (1-2431)

Qy 1 MetanlyspProilethProserThrTyValArgCysLeuAanValGlyLeuileArg 20
 Db 168 ATGAACAAGCCGTTGACACCATCGACATACATACGCAACCTTAATGTGGGATCTTAAG 227
 Qy 21 LysLeuSeraspPheileaspProGlnGluGlyTyrPlySylsLeuAlaValAlaIlelys 40
 Db 228 AAGCTGCGATTTATTTGATCTTCACAGAGGGTGAAGAATTAGACGTGATCAAA 287
 Qy 41 LysProserGlyAaspAaspArgTyrAsnGlnPheHisIleArgAargPheGluAlaLeuLeu 60
 Db 288 AAGCGGTCGGCGACGACATACATCAATTCATACAGAGATTCGAAAGCCCTTACTT 347
 Qy 61 GlnThrGlyLysSerProThrsSerGlyLeuLeuPheAspTyrPglYThrThrAsnCysThr 80
 Db 348 CAGACCGGAGAGACCCCACTGTGAATCTGCTTTGACTGGGACACCAACATCGACA 407
 Qy 81 ValGlyAaspLeuValAaspLeuLeuileGlnAsnGlnPheAlaProAlaSerLeuLeu 100
 Db 408 GTTGCGCATCTTGATGATCTACTGGTCCAGATTGAGCTGTTGCCCGCCGACCTCTCTG 467
 Qy 101 LeuProaspAlaValProlyThrAlaAsnThrLeuProserLysGluAlaIleThrVal 120
 Db 468 CTGCGGATGCGCTTCCCAACCGTCAAAAGCCCTGCTCTAGAGAGCGGCAACAGTG 527
 Qy 121 GlnGlnLysGlnMetProPheCysAaspLysAaspArgThrLeuMetThrProValGlnAsn 140
 Db 528 GCACAAACACCGGCGCTTTCAGAGAAAGACAGACATCCGTAATCCCTAATGCCAG 587
 Qy 141 LeuGlnLysSerTyrMetProProAaspSerSerProGlnLysSerLeuGluVal 160
 Db 588 CTAGAACACACGCTGCGAGACCGGACTCTCAAGCCCAACAGAGGTAGAGTCC 647
 Qy 161 SeraspThrArgPheHisSerPheSerPheTyrGlyLeuLysAanValThrAsnAspPhe 180
 Db 648 AGCACACTCGGTTCCACAGCTTCTCGTTCCATGATGAGAGCATCAACAACTCTC 707
 Qy 181 AspGlnLysProIleSerValGlyLysAsnMetGlyGluGlyPheGlyValVal 200
 Db 708 GACACGACACCCGGCTCGCCGGTGGCAACCGATGGAGAGGGGATTTGGATGTG 767
 Qy 201 TyrLysGlyTyrValAsnAsnThrThrValAlaValLysLysLeuAlaIleMetValAsp 220
 Db 768 TACAAGGGCTGTGTGAACACACCATCTGGCGGTGAAGAAGCTCGAGCATGTGTGA 827
 Qy 221 IleThrThrGlnGluLeuLysGlnGlnPheAspGlnGluIleLysValMetAlaLysPhe 240
 Db 828 ATCAAGTACTGAAGACTTAAGCAACAGTTGATCAAGAAATTAAGTATGCAACGCT 887
 Qy 241 GlnHisGluAanLeuValGlyLeuLeuLysPheSerSerAspGlyAaspAspLeuCysLeu 260
 Db 888 CAGCAGCAGAACTGGTGAAGCTGCTCGGCTTCTCCAGCAGACAGGACACATCTGTCT 947
 Qy 261 ValTyrValTyrMetProAsnGlySerLeuLeuAaspArgLeuSerCysLeuAspGlyThr 280
 Db 948 GTGATGTCTTACATGCCCAAGGGCTTGTGCGGACACATCTGCTGCTGATGTGATCA 1007
 Qy 281 ProProLeuSerTyrPheMetArgCysLysIleAlaGlnGlyAlaAlaAsnGlyIleAsn 300
 Db 1008 CCACCGCTTCTCTGCGACACACAGGTGCAAGGTCTCGAGGACACGCAAAATGCGATCGG 1067
 Qy 301 PheLeuHisGluAanHisIleIleHisArgAspIleLysSerAlaAsnIleLeuLeuAsp 320
 Db 1068 TTCTGCAATGAATCATCACTCATATGAGATTAATAAGGCAAAATATCTACTAGAC 1127
 Qy 321 GluAlaPheThrAlaLysIleSerAaspPheGlyLeuAlaArgAlaSerGluLysPheAla 340

Db 1128 AGACACTTACTCCAAAATATCGACTTGGGCTTGCACGGGCTTCGCAAGGCTAGCG 1187
 Qy 341 GlnThrValMetThrSerArgIleValGlyThrThrAlaTyrMetAlaProGluAlaLeu 360
 Db 1188 CAGACGCTCATGACACCGCAATCTGGGACACACCGCTTACATGACCGCAACGCTTGG 1247
 Qy 361 ArgGlyGluIleThrProLysSerAspIleTyrSerPheGlyValValLeuLeuGluIle 380
 Db 1248 CCGGAGAAATTAACCCCAATCTGACATCTACAGCTTGGCGGTGTTGTTGAGCTG 1307
 Qy 381 IleThrGlyLeuProAlaValAaspGluHisArgGluProGlnLeuLeuLeuAspIleLys 400
 Db 1308 ATAAACCGGCTGCGGCTGATGATGAACCGGAACCTCACTACTGCTGATATTAA 1367
 Qy 401 GlnGluIleGlnAaspGlnGluLysThrIleGlnAaspTyrIleAaspLysMetAsnAsp 420
 Db 1368 GAAAGATTAAAGTGAAGAGAGACATTTGAAAGTTTACCGATGAGAGATGAGCAT 1427
 Qy 421 AlaAaspSerThrSerValGluAlaMetTyrSerValAlaSerGlnCysLeuHisGluLys 440
 Db 1428 GCGGACCTGCTTCCGTTGGAAGCAATGATCTGCTGCTAGCCAGTGTCTGATGAGAAG 1487
 Qy 441 LysAanLysArgProAaspIleLysLysValGlnGlnLeuLeuGlnGlnMetThrAla 459
 Db 1488 AAAAACAAGCGCCAGACATTCGAAAGTTCAACAGCTGCTACAGAGATGTCTCT 1544

RESULT 11
 LOCUS BD155790 2213 bp DNA linear PAT 17-JAN-2003
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BD155790
 VERSION BD155790.1 GI:27861548
 KEYWORDS JP 2002191363-A/10633.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2213)
 Oca,T., Isogai,T., Nishikawa,T., Hayaishi,K., Saito,K., Yamamoto,D.,
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 10633 09-JUL-2002;
 JOURNAL HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/10633
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUJI OTSUKI
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10,
 PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 Location/Qualifiers
 FT CDS (977).. (1864).
 FEATURES
 source Location/Qualifiers
 1..2213
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 642 a 488 c 516 g 567 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 9,7e-146 Length: 2213
 Scores: 1787.50 Matches: 352
 Percent Similarity: 87.34% Conservative: 0
 Best Local Similarity: 87.34% Mismatches: 1
 Query Match: 74.92% Indels: 50

DB: 6 Gaps: 1
 US-10-001-254-16 (1-460) x BD155790 (1-2213)

QY 1 MetAnlyseProleTherProserThyThraValaGySleuAsnValGlyLeuIleArg 20
 Db 751 ATGAACCAACCCATACACACATCAATATGCGCCCTGCTCAAGTGTGACCTAAATTAGG 810

QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTTPVLysLeuAlaValAlaIleLys 40
 Db 811 AAGCTGATGATTTTATTCATCTCTCAAGAGAGATGAGAGATGAGCTGTGATCTATTA 870

QY 41 LysProSerGlyAspAspArgTyraGlnGlnPheHisIleArgArgPheGluAlaLeu 60
 Db 871 AAACCATCTGGTATGATGATACATCACTTCAATAGAG----- 912

QY 61 GlnThrGlyLysSerProTherSerGluLeuLeuPheAspTTPGlyThThraAsnCysThr 80
 Db 912 ----- 912

QY 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPhePheAlaProAlaSerLeu 100
 Db 912 ----- 912

QY 101 LeuProAspAla-ValProLysThraAlaAsnThreLeuProSerLysGluAlaIleThVa 120
 Db 913 -----TGCTGTTCCCAAAACGCTAATACATCACTTCTTAAGAGAGCTAATACAGT 963

QY 120 GlnGlnLysGlnMetProPheCysAspLysAspArgThreLeuMetThrProValGlnAs 140
 Db 964 TCAGCAAAAACAGATGCTTCTCTGACAAAGACAGACATGATGACACCTGTGCAGAA 1023

QY 140 nLeuGlnGlnSerThyMetProProAspSerSerProGluAsnLysSerLeuGluVa 160
 Db 1024 TCTTGAACCAAGCTATATGCACTGCTCCCAAGTCCAGAAATTAAGATTGAAGGT 1083

QY 160 LserAspThraArgPheHisSerPheSerPheTyrgLysLeuLysAsnValThraAsnAsp 180
 Db 1084 TAGGATATGACAGTTTTCACAGTTTTCATTTTATTAATGAAGATGTACAAATATACCT 1143

QY 180 eAspGluArgProIleSerValGlyLysAsnLysMetGlyGluGlyPheGlyValVa 200
 Db 1144 TGATGACAGACCAATTTCTGTTGCTGATTAATTAATGAGAGAGAGATTTGAGTGT 1203

QY 200 LThyLeuGlyThyThraLysAsnThrThraValaValLysLysLeuAlaIleMetValAs 220
 Db 1204 ATATTAAGGCTACGTAATTAACACACATGCGACGTGAAGAGCTTGAAGCAATGTGTA 1263

QY 220 PileThrThrgLysLysGlnGlnPheAspGlnGlnLysValMetAlaLysCys 240
 Db 1264 CATTACTACTGAGAGAACTGAAACAGCACTTTCATCAAGAAATTAATATGCGAAATG 1323

QY 240 sGlnHisGluLysLeuValGluLeuLeuGlyPheSerSerAspGlyAspAspLeuCysLe 260
 Db 1324 TCAACATCAAAAACCTAGTAGAACTACTGTTCTTCAAGTATGAGATGACCTGCTT 1383

QY 260 uValThyThraLysMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyTh 280
 Db 1384 AGATATGTTTACATGCTTAAAGGTTCAATGCTTACACAGACTCTCTTCTTGAGATG 1443

QY 280 rProProLeuSerThraPheMetArgCysLysIleAlaGlnGlnValAlaAsnGlyLys 300
 Db 1444 TCCACCACTTCTTGGCAGATGAGATGCAAGTTCTCAGGGTGAAGCTAATGGATCA 1503

QY 300 nPheLeuHisGluLysAsnHisIleHisArgAspLysSerAlaAsnIleLeuLys 320
 Db 1504 TTTTCTACATGAATAATCATATATCATAGATTTAAAGTGAATAATATCTTACTGGA 1563

QY 320 PGLuAlaPheThraLysLysIleSerAspPheGlyLeuAlaArgAlaSerGlyLysPheAl 340
 Db 1564 TGAAGCTTTTACTGTAATAATATCTGACTTGTGCGCTTGTGAGAAATTTGCT 1623

QY 340 aGlnThraValMetThraSerArgIleValGlyThThraLysMetAlaProGluAlaLe 360

Db 1624 CCAGCACTCATGATGATGACAGAAATTTGTGGAACAACAGCTTATATGACCAAGAGCTTT 1683

QY 360 uArgGlyGluIleThrProLysSerAspLysIleTySerPheGlyValLeuLeuGluIle 380
 Db 1684 GCGTGGAGAAATAACACCAATCTGATATTTACAGCTTGTGTGTGTTTACTAGAAAT 1743

QY 380 eIleThrGlyLeuProAlaValAspGluHisArgGluProGlnLeuLeuAspLysLe 400
 Db 1744 AATTACTGACTTCCAGCTGTGATGAGACACCGTGAACCTCAGTTATTCATGATATTA 1803

QY 400 sGluGlu 402
 Db 1804 AGAAGA 1810

RESULT 12
 AK027301 2213 bp mRNA linear PRI 01-AUG-2002
 LOCUS Homo sapiens CDNA FLJ14395 file, clone HEMBA1003250, weakly similar
 DEFINITION to PROTEIN KINASE APKIA (BC 2.7.1.-).
 AK027301
 ACCESSION AK027301.1 GI:14041890
 VERSION AK027301.1 GI:14041890
 KEYWORDS oligo capping, fls (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Isogai,T., Oca,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Wagatsuma,M., Hosokiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuo,Y., Nishimura,K. and Iwayanagi,T.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2213)
 AUTHORS Isogai,T. and Otsuki,T.
 TITLE Direct Submission
 JOURNAL Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
 FEATURES
 SOURCE
 1. 2213
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HEMBA1003250"
 /tissue_type="whole embryo, mainly head"
 /clone_id="HEMBA1"
 /dev_stage="embryo, 10 weeks"
 /note="cloning vector: pMB18SFL3"

BASE COUNT 642 a 488 c 516 g 567 t
 ORIGIN

Alignment Scores:
 Pred. No.: 9,7e-146 Length: 2213
 Score: 1787.50 Matches: 352
 Percent Similarity: 87.34% Conservative: 0
 Best Local Similarity: 87.34% Mismatches: 1
 Query Match: 74.92% Indels: 50
 DB: 9 Gaps: 1

US-10-001-254-16 (1-460) x AK027301 (1-2213)

QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
 Db ATGAAACAAACCCATTAACACCATCAACATGTCGCTCCATGTTGACATAATTAGG 810
 QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTTPlyLysLeuAlaValAlaIleLys 40
 Db AACCTGCAATTTTATTTGATCTCCAAAGAAAGAAAGAAAGTTTAAAGCTGATTTTAA 870
 QY 41 LysProSerGlyAspAspArgTyraAsnGlnPheHisIleArgArgPheGluAlaLeuLeu 60
 Db AAACCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 912
 QY 61 GlnThrGlyLysSerProThrSerGluLeuLeuPheAspTrpGlyThrThrAsnCysThr 80
 Db 912 ----- 912
 QY 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPheAlaProAlaSerLeuLeu 100
 Db 912 ----- 912
 QY 101 LeuProAspAla-ValProLysThrAlaAsnThrLeuProSerLysGluAlaIleThrVa 120
 Db 913 -----TGCCTTCCCAAACTGCTTAATACCTACTTTTAAAGAGCTATTAACAGT 963
 QY 120 GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAs 140
 Db TCAGCAAAAACAGATGCTTCTTCTGACAAAGACAGACATGATGACACTGTCAGCAA 1023
 QY 140 LysGlnGlnSerTyMetProProAspSerSerSerProGlnAsnLysSerLeuGluVal 160
 Db 1024 TCTTGAACAAAGCTATATGACCACTGACTCTCAAGTCCAGAAATTAAGTTTGAAGT 1083
 QY 160 LysAspThrArgPheHisSerPheSerPheTyGlyLeuLysAsnValThrAsnAsnPh 180
 Db 1084 TAGGATACACGCTTTTACAGCTTTTCACTTTTATGATGATGATGATGATGATGATGAT 1143
 QY 180 AspGluArgProIleSerValGlyGlyAsnLysMetGlyGlyGlyGlyPheGlyValVa 200
 Db 1144 TGATGAACGACCCATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1203
 QY 200 LysTyrLysGlyTyrValAsnAsnThrThrValAlaValLysLeuAlaIleMetValAs 220
 Db 1204 ATATTAAGGCTACCTTAATAACACACTGTGGCAGTAAAGCTTCCAGCAAGTATTAAGT 1263
 QY 220 PileThrThrGluGluLeuLysGlnGlnPheAspGlnGluIleLysValMetAlaLysCy 240
 Db 1264 CATTACTACTGAGAACTGAAACGCGCTTGTATCAAGAAATTAAGTAATGCGAAAGTG 1323
 QY 240 SglnHisGluAsnLeuValGluLeuLeuGlyPheSerSerAspGlyAspAspLeuCysLe 260
 Db 1324 TCAAACATGAANAATTAGTAGAATACTGTTGTTCTCAAGTAGATGAGATGACCTGCTT 1383
 QY 260 ValTyrValTyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyTh 280
 Db 1384 AGATATAGTTTACATGCTTAACGCTTCACTGCTGAGAACAGACTCTGCTGATGATGAT 1443
 QY 280 rProProLeuSerTrpHisMetArgCysLysIleAlaGlnGlyValAlaAsnGlyLys 300
 Db 1444 TCCACCACTTTCTTGGCAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1503
 QY 300 nPheLeuHisGluAsnHisIleHisIleArgAspIleLysSerAlaAsnIleLeuLeuAs 320
 Db 1504 TTTTCTCATGTAATAATCATCATTAATCAAGATATTAAGTAATTAATTAATTAATTA 1563
 QY 320 PGIuAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGlyLysPheAl 340
 Db 1564 TGAAGCTTTTACTGCTAAATATCTGACTTTGGCTTGCACGGCTTCTGAGAAAGTTTGC 1623
 QY 340 agInThrValMetThrSerArgIleValGlyThrThrAlaTyrMetLysProGluAlaLe 360
 Db 1624 CCGACACGTATGATGACAGAAATTTGGGAAACACGCTTATATGACACCAAGAGCTTT 1683
 QY 360 uArgGlyGluIleThrProLysSerAspIleTyrSerPheGlyValValLeuLeuGluI 380

Db 1684 GCGTGGAGAAATTAACCCCAATCTGATATTAACACTTGGTGTGTTTACTAGAAAT 1743
 QY 380 eIleThrGlyLeuProAlaValAspGluHisArgGluProGlnLeuLeuAspIleLys 400
 Db 1744 AATTAAGTGGACTTCCAGCTGTGATGAACACCGTGAACCTCACTTATTTGCTAGATTTAA 1803
 QY 400 sGluGlu 402
 Db 1804 AGAAGAA 1810
 RESULT 13
 AR223870 833 bp DNA linear PAT 26-SEP-2002
 LOCUS Sequence 10 from patent US 6440663.
 DEFINITION AR223870
 ACCESSION AR223870.1 GI:23332452
 VERSION AR223870.1 GI:23332452
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 833)
 AUTHORS Seanlan,M.J., Stockert,E., Chen,Y.-T., Old,L.J., Jager,E. and Knuth,A.
 JOURNAL Renal cancer associated antigens and uses therefor
 FEATURES Patent: US 6440663-A 10 27-AUG-2002;
 source Location/Qualifiers
 1..833
 BASE COUNT 273 a 155 c 179 g 226 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,56e-92 Length: 833
 Score: 1169.00 Matches: 246
 Percent Similarity: 93.21% Conservative: 1
 Best Local Similarity: 92.83% Mismatches: 8
 Query Match: 48.99% Indels: 10
 DB: 6 Gaps: 3
 US-10-001-254-16 (1-460) x AR223870 (1-833)
 QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
 Db 50 ATGAAACAAACCCATTAACACCATCAACATGTCGCTCCATGTTGACATAATTAGG 109
 QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTTPlyLysLeuAlaValAlaIleLys 40
 Db 110 AACCTGCAATTTTATTTGATCTCCAAAGAAAGAAAGAAAGTTTAAAGCTGATTTTAA 169
 QY 41 LysProSerGlyAspAspArgTyraAsnGlnPheHisIleArgArgPheGluAlaLeuLeu 60
 Db 170 AAACCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 229
 QY 61 GlnThrGlyLysSerProThrSerGluLeuLeuPheAspTrpGlyThrThrAsnCysThr 80
 Db 230 CAAACATGTAATAATCTGCTTAATCACTGTTGATGATGATGATGATGATGATGATGAT 289
 QY 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPheAlaProAlaSerLeuLeu 100
 Db 290 GTTGGATGATCTTGTGATCTTGTGATCCAAAGAAATTTTGTCTCTGAGATCTTTTG 349
 QY 101 LeuProAspAlaValProLysThrAlaAsnThrLeuProSerLysGluAlaIleThrVal 120
 Db 350 CTCCTCGATGCTTCTTCCAAACCTGCTTAATACCTACTTTTAAAGAGCTATTAACAGTT 409
 QY 121 GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140
 Db 410 CAGCAAAAACAGATGCTTCTTCTGACAAAGACAGACATGATGACACTGTCAGCAAGAT 469
 QY 141 LeuGlnGlnSerTyMetProProAspSerSerSerProGlnAsnLysSerLeuGluVal 160
 Db 470 CTGGAACAAAGCTATATGACCACTGACTCTCAAGTCCAGAAATTAAGTTTGAAGTT 529

QY 161 SerAspThrArgPheHisSerPheSerPheTyrgluLeuValAsnValThrAsnAspHe 180
DB 530 AGTGATACACGTTTTCACAGCTTTTTCATTTATGATTGAGAAATGTCAAAATACCTT 589
QY 181 AAspGluArgProIleSerValGlyValMetGlyGlyGlyPheGlyValVal 200
DB 590 GATGACACACCCATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 649
QY 201 Tyr-LysGlyTyrValAsnAsnThrThrValAlaVal-LysLysLeuAlaIleMetValA 220
DB 650 TATAAAGGCTTACGTAATTAACAACTGTGCACTGAAAGAAAGCTTGCAGCAATGCTTG 709
QY 220 spIIeThrThrGlu-GluLeuLys-GlnGlnPheAsp-GlnGlu-IleLysValMetAla 238
DB 710 ACATTACTACTGAAAGAACTGAAACAGCAGTTTGTATCCAAAGAAATAAATATGAGC 769
QY 239 LysCysGlnHis-----GluAsnLeuVal---GluLeuLeuGlyPheSerSer---Asp 254
DB 770 AAAAGTGTCAACATGCAAAAACTTGTAGTAAAGAACTTGTGTGTGTGTGTGTGTGTGT 829
QY 255 Gly 255
DB 830 GGG 832

RESULT 14
BC045381 1719 bp mRNA linear VRT 10-FEB-2003
LOCUS Dantio rerio, similar to interleukin-1 receptor-associated kinase 4,
DEFINITION clone MGC:55553 IMAGE:2642773, mRNA, complete cds.
ACCESSION BC045381
VERSION BC045381.1 GI:28278875
KEYWORDS MGC.
SOURCE Dantio rerio (zebrafish)
ORGANISM Dantio rerio

REFERENCE
AUTHORS Straubeberg, R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
REMARK Contact: MGC help desk
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami
cDNA Library Preparation: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
Contact: amadansystemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whitting

FEATURES
SOURCE
1..1719
/organism="Dantio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:55553 IMAGE:2642773"
/rissue_type="whole body, adult, (one male and one female,
including unfertilized eggs)"
/clone_lib="Sugano_Kawakami_zebrafish_DRA"
/lab_host="DH10B"

CDS
/note="Vector: pME18S-FL3"
122..1270
/codon_start=1
/product="similar to interleukin-1 receptor-associated
kinase 4"
/protein_id="AA045381.1"
/db_xref="gi:28278876"
/translation="MSDVTTPNTPRKLRYSALRALADLDPODTWRSIMADISRPCGE
PRYOMHRREPAACVLOGKSPYMLLEPDMGTSCDGLVLEILRHOLFVAVTLPLP
HSVCHAHGVSVCSEAAPASAVICLACEITOTVDNSNKPQKISKPVEDDSNKPQEO
ELFVPRPSSSGAOSSSMDSOGFRTPSHELTANTQWDERPLSDGCRLOSGRGVY
FRGMWDHVAVKLNPIDSSYEDLKQFNOELOTLSLSHENVALRVLGSCGPPPL
CVPELWNGSLERLACERTPALTRNRCMTITVGARGSLYATHVAHHRDVKSAN
ILDEGFVAKISDFGLTRSAAGAAGFVWASSGRSGSCTVADGEG"

BASE COUNT 492 a 399 c 477 g 351 t
ORIGIN

Alignment Scores:
Pred. No.: 8.82e-76 Length: 1719
Score: 981.00 Matches: 199
Percent Similarity: 61.34% Conservative: 93
Best Local Similarity: 41.81% Mismatches: 121
Query Match: 41.11% Indels: 63
DE: 5 Gaps: 4

US-10-001-254-16 (1-460) x BC045381 (1-1719)

QY 5 IleThrProSerThrTyrValArgCysLeuAsnValGlyLeuIleArgLysLeuSerAsp 24
DB 131 GTCCAGCGCGAACAACCGCGTGTGAAGCTCCGGTACTCCGCTAGAGCTTTGACAC 190
QY 25 PheIleAspProGlnGluGlyTyrLysLeuValAlaIleLysLysProSerGly 44
DB 191 CTGCTGACCCCGACAGACCTTGAGAGCATCTGCGGACATCAGACACCTTCCGG 250
QY 45 AspAspArgTyrAsnGlnPheHisIleArgArgPheGluAlaLeuLeuGlnThrGlyLys 64
DB 251 GAACCCAGATACACACAGATGACATGAGCGGCTTGAAGCATCGCTCCAGAGAAAG 310
QY 65 SerProThrSerGluLeuLeuPheAspTyrGlyThrThrAsnGlyThrValGlyAspLeu 84
DB 311 AGTCCACACATGACATGCTGCTGTGACCTGGGGAACCTCAGACTGCTGTTGAGATCTG 370
QY 85 ValAspLeuLeuIleGlnAsnGluPhePheAlaProIleAspLeuLeuProAsp--- 103
DB 371 GTGAGATCTTGATCCGATCAGCTGTCGCCCGCTCAGATCTGCTGCTGCGATAC 430
QY 104 -----AlaValProLysThrAla 109
DB 431 AGCGTGTGACACGCTCAACAGGCTCAGTGTGCTGAGAGAGGCGCTCCGCGTGTGCA 490
QY 110 AsnThrLeuProSerLysGluAlaIleThrValGlnGlnLysGlnMetProPheCysAsp 129
DB 491 GTGTGTCTG-----CAGGCACTGCAAAATCACCCAACTGTAGAGACAAACATTAAT 541
QY 130 LysAspArgThrIleuMetThrProValGlnAsn-----LeuGlu 142
DB 542 AAACCCCAAAAATCTCTAAACCTGTGAGAGATACAGTAATAAACCCGTCACAGACAG 601
QY 143 GlnSerTyrMetProProAspSerSerSerProGluLeuLysSerLeuGluValSerAsp 162
DB 602 GAGCTGTTTGTGAGCCGAGACAGACTCTGAGACACAGAGAGACACTGGATCTCCAGC 661
QY 163 ThrArgPheHisSerPheSerPheTyrgluLeuLysAsnValThrAsnAspAspGlu 182
DB 662 CAAGGCTTCCACACCTTCAGCTTCAGTGAAGTGAAGCGCATGACGACGACGAGAGAG 721
QY 183 ArgProIleSerValGlyValMetGlyGlyGlyPheGlyValValTyrLys 202
DB 722 AGGCGACTTTCGATGAGGCGCTGCGGCTGAGGCTTGAAGGCTTGTAGAG 781
QY 203 GlyTyrValAsnAsnThrThrValAlaValLysLysLeuAlaIleMetValAspIleThr 222

```

Db      782  GGACGCGTGGGGGTAACATGTGCGGTCAGAAACTCAATCTTGGACGGCAGTTCA 841
Qy      223  ThrGluLeuLeuysGlnGlnPheAspGlnGluLeuValMetAlaLysCysGlnHis 242
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      842  TATGAAAGACCTCAAGAGAGTTTAAACAGAGATTGACACATCAGAAAGTCATGATCAT 901
Qy      243  GluAsnLeuValGluLeuLeuGlyPheSerSerAspGlyAspAlaLeuValTyr 262
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      902  GAGAAATGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 961
Qy      263  ValTyrMetProAsnGlySerLeuLeuAspArgPheSerCysLeuAspGlyThrPro 282
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      962  GAGCTGATGTGTAACGGCTCTCTCTGAGCGACTGCGCTGCTGTAACACACACTGGC 1021
Qy      283  LeuSerTrpHisMetArgCysLysLeuAlaGlnGlyAlaAlaAsnGlyLeuAsnPhe 302
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1022  CTCACCTGAGAGAACATATGTGATCATCTGTGGAGTGTGCCGGGCTGAGCTTACTG 1081
Qy      303  HisGluAsnHisHisLeuArgAspLeuLysSerAlaAsnLeuLeuAspGluAla 322
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1082  CACACACACGACATCCACAGAGAGCTCAAGAGCCCAACATCTTAAGACGAGGCT 1141
Qy      323  PheThrAlaLysHisSerAspPheGlyLeuAlaArgAlaSerGlyLysPheAlaGln 342
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1142  TTGTGTGCGAGAGATCTCAGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1177
Qy      343  ValMetThrSerArgLeuValGlyThrThrAlaTyrMetAlaProGluAlaLeuArg 362
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1177  ----- 1177
Qy      363  GluIleThrProLysSerAspIleTyrSerPheGlyVal-ValLeuLeuGluIleIle 382
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1178  -----TCTGCGGCTGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1210
Qy      382  rGlyLeuProAlaValAspGluHisArgGluProGlnLeuLeuLeuAspIleLysGlu 402
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1211  TGGGCTTCTCCACGTGAGCGAGAGTGGGATCTGCATCTGTGAGGAGAAAGATGA 1270
Qy      402  uIleGluAspGlnGluLysThrIleGluAspTyrIleAspLysLysMetAsnAlaAs 422
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1271  TCTGAGATGAGAGATCTCTCTGCTGAGACTTCACTGACGCCGCGCATGCGAGCTG 1330
Qy      422  pSerThrSerValGluAlaMetTyrSerValAlaSerGlnCysLeuHisGlyLysLys 442
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1331  AACAGAGAGACTGACATGATGTAAGCCGCTCTCAGTGTGTGCCAGAGAGAA 1390
Qy      442  nLysArgProAspIleLysLysValGlnGlnLeuLeuGlnIleMet 457
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1391  CAAGAGACCCGCCATGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1436

```

```

RESULT 15
AX053431          539 bp   DNA      linear   PAT 13-JAN-2001
LOCUS             AX053431
DEFINITION        Sequence 197 from Patent WO0073801.
ACCESSION         AX053431
VERSION           AX053431.1 GI:12227750
KEYWORDS
SOURCE
ORGANISM          Homo sapiens (human)
                  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS           Obata, Y.
TITLE             Breast, gastric and prostate cancer associated antigens and uses
                  thereof
JOURNAL           Patent: WO 0073801-A 197 07-DEC-2000;
                  LUDWIG INSTITUTE FOR CANCER RESEARCH (US)
FEATURES
Source            1..539
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"

```

```

BASE COUNT      168 a 94 c 111 g 163 t 3 others

```

```

ORIGIN
Alignment Scores:
Pred. No.:      6,4e-68
Score:          883.00
Percent Similarity: 97.24%
Best Local Similarity: 97.24%
Query Match:    37.01%
DB:             6
Gaps:           0
US-10-001-254-16 (1-460) x AX053431 (1-539)

```

```

Qy      258  LeuCysLeuValTyrValTyrMetProAsnGlySerLeuLeuAspAlaLeuSerCysLeu 277
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1  CTCGCTTATATATATGTTTACATGCCCTAATGATGTTCAATGCTTACAGACACTCTTGC 60
Qy      278  AspGlyThrProProLeuSerTrpHisMetArgCysGlyLeuAlaGlnGlyAlaAsn 297
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      61  GATGTACTCCACACACTTCTTGTGACATGATGCAAGATGCTCGGGGTGACGTAAT 120
Qy      298  GlyIleAsnPheLeuHisGluAsnHisHisLeuArgAspIleLysSerAlaAsnIle 317
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      121  GGCATCAATTTCTTACATGAAATATCATCATATTCATAGAGATTTAAAGTCAAAATATC 180
Qy      318  LeuLeuAspGluAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGlu 337
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      181  TTACTGATGATAGCTTTTACTGCTAAATATCTGACTTTGGCTTGACGGGCTTCTGAG 240
Qy      338  LysPheAlaGlnThrValMetThrSerArgIleValGlyThrThrAlaTyrMetAlaPro 357
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      241  AAGTTTGCCAGACAGTCACTAGCAGATGTGGAAACAGCTTATATGACCA 300
Qy      358  GluAlaLeuArgGlyGluIleThrProLysSerAspIleTyrSerPheGlyValIleLeu 377
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      301  GAAGCTTGGGTGAGAAATTAACACCAATCGATATTACAGCTTGTGTGTTTA 360
Qy      378  LeuGluIleIleThrGlyLeuProAlaValAspGluHisArgGluProGlnLeuLeu 397
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      361  CTGAATATATTAATCTGACTTCCAGCTGTGATGAACACCGGGAACCTCAGTTATTGCTA 420
Qy      398  AspIle-LysGluGluIleGluAspGluLysThrIleGluAspTyrIleAspLysGly 417
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      421  GATATTAAAGAAATTTGNAATGAGAAAGACATTTGAAGATTATTAATAAAAA 480
Qy      417  smetAsnAspAlaAspSerThrSerValGluAlaMetTyr-SerValAlaSerGlnCys 436
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      481  GATGAATGATGCTGATTCCTCAGTTGAGTGAANCTATGTACTCTGTGTAATGATCAATGT 539

```

```

Search completed: January 18, 2004, 04:26:40
Job time : 7811.21 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 18, 2004, 01:16:08 ; Search time 120.032 seconds
(without alignments)
1691.512 Million cell updates/sec

Title: US-10-001-254-16
Perfect score: 2386
Sequence: 1 MNKPITPSTYVRCINLVGHIR.....KKRPDIKKVQQLQEMTAS 460

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 segs, 220691566 residues
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+.p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10001254/runat.16012004.152425.19740/app_query.fasta.1.1109
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rm1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10001254 @CGN 1.1.95 @runat.16012004.152425.19740 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Issued Patents_NA:*
- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
 - 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
 - 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
 - 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
 - 5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
 - 6: /cgn2_6/ptodata/1/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1169	49.0	833	4	US-09-166-350-10
2	515.5	21.6	3590	4	US-08-587-889-1
3	515.5	21.6	3590	4	US-09-016-434-1093
4	515.5	21.6	3590	5	PCT-US96-09193-1
5	476.5	20.0	1488	4	US-09-579-182-4
6	452.5	19.0	2336	4	US-09-228-986-10
7	451	18.9	3239	4	US-09-228-986-9
8	448.5	18.8	2432	4	US-09-228-986-7
9	438	18.4	2288	3	US-09-135-232-1
10	438	18.4	2288	4	US-09-863-549-1
11	432	18.1	2868	4	US-09-228-986-4
12	430.5	18.0	966	1	US-08-447-185-2

13	430.5	18.0	2443	1	US-08-447-185-3	Sequence 3, Appl1
14	427.5	17.9	2749	1	US-07-717-331F-4	Sequence 4, Appl1
15	424	17.8	3459	2	US-08-980-060-3	Sequence 3, Appl1
16	424	17.8	3459	3	US-09-307-185-3	Sequence 3, Appl1
17	423	17.7	1806	2	US-08-980-060-1	Sequence 1, Appl1
18	423	17.7	1806	3	US-09-307-185-1	Sequence 1, Appl1
19	410.5	17.2	2943	4	US-09-503-922-2	Sequence 2, Appl1
20	403	16.9	283	4	US-09-016-434-998	Sequence 998, App
21	402.5	16.9	2749	1	US-08-265-678-1	Sequence 1, Appl1
22	398.5	16.7	2571	1	US-07-717-331F-9	Sequence 9, Appl1
23	398.5	16.7	2833	1	US-07-717-331F-1	Sequence 1, Appl1
24	391.5	16.4	1029	4	US-09-228-986-11	Sequence 11, Appl1
25	390.5	16.4	1294	4	US-09-579-182-3	Sequence 3, Appl1
26	386	16.2	1554	2	US-08-587-860A-24	Sequence 24, Appl1
27	386	16.2	4104	3	US-08-881-706-1	Sequence 1, Appl1
28	382.5	16.0	2389	4	US-09-228-986-1	Sequence 1, Appl1
29	379	15.9	5733	2	US-08-473-553A-1	Sequence 1, Appl1
30	377	15.8	2513	4	US-09-228-986-13	Sequence 13, Appl1
31	358.5	15.0	2686	4	US-09-228-986-3	Sequence 3, Appl1
32	341.5	14.3	2510	2	US-08-888-982A-42	Sequence 42, Appl1
33	341.5	14.3	2510	4	US-09-462-261-42	Sequence 42, Appl1
34	341.5	14.3	2510	4	US-09-506-073-89	Sequence 89, Appl1
35	337.5	14.1	2638	4	US-09-228-986-8	Sequence 8, Appl1
36	321	13.5	1947	1	US-08-185-282-11	Sequence 11, Appl1
37	321	13.5	2977	1	US-08-276-151-1	Sequence 1, Appl1
38	321	13.5	2977	1	US-08-306-691B-17	Sequence 17, Appl1
39	321	13.5	2977	2	US-08-756-806A-64	Sequence 64, Appl1
40	321	13.5	2977	3	US-08-328-239A-4	Sequence 4, Appl1
41	321	13.5	2977	3	US-09-143-214-64	Sequence 64, Appl1
42	321	13.5	2977	3	US-09-209-668-12	Sequence 12, Appl1
43	321	13.5	2977	4	US-09-506-073-25	Sequence 25, Appl1
44	321	13.5	2977	4	US-09-555-279-1	Sequence 1, Appl1
45	321	13.5	2977	5	PCT-US95-13661-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-166-350-10
; Sequence 10, Application US/09166350A
; Patent No. 6440663
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Chen, Yao
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Old, Lloyd
; APPLICANT: Jager, Alex
; TITLE OF INVENTION: Renal Cancer Associated Antigens and
; TITLE OF INVENTION: Uses Therefor
; FILE REFERENCE: L0461/7051
; CURRENT APPLICATION NUMBER: US/09/166,350A
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: US 09/166,350
; EARLIER FILING DATE: 1998-10-05
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 833
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-166-350-10
Alignment Scores:
Pred. No.: 1,31e-127 Length: 833
Score: 1169.00 Matches: 246
Percent Similarity: 93.21% Conservative: 1
Best Local Similarity: 92.83% Mismatches: 8
Query Match: 48.99% Indels: 10
DB: 4 Gaps: 3
US-10-001-254-16 (1-460) x US-09-166-350-10 (1-833)

Qy		MetAsnLysProIleThrProSerThrTyrValA-GCySLeuAsnValGlyLeuIleArg	20
Db		ATGAAACAAACCCATAACCATACATATGAGCGCTGCATATGTTGACTATTAAGG	1089
Qy		LysLeuSerSerPheIleAspProGlnGluGlyTyrPlySlyLeuValaValaIleLys	40
Db		110 AAGGTGTCAGATTTTATTTATGATCTCTCAAGAAAGATGGAAGAGTTAGCTGTAGCTATTTAAA	1694
Qy		LysProSerGlyAspAspArgTyrAsnGlnPheHisIleLeuArgPheGluValaLeuLeu	60
Db		170 AAACCACTGTGTATGATATGATATCAATCAGTTTCAACATAAGAGATTTTGAAGCACTTACTT	2238
Qy		GlyThrGlyLysSerProThrSerGlnLeuLeuPheAspTyrGlyThrThrAsnCySThr	80
Db		230 CAACCTGGAAAAAGCTCCCACTTGTGAATTTACTGTTTGACTGGGGACACACAAATTTGACA	2888
Qy		ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPhePheAlaProAlaSerLeuLeu	1008
Db		290 GTTGATGATCTGTGGATCTTTTGGATCCAAATATGAATTTTTTGGCTCGCGAGCTTTTG	3498
Qy		LeuProAspAlaValaProLysThrAlaAsnThrLeuProSerLysGluValaIleThrVal	1208
Db		350 CTCACAGATCTGTCTCCCAAAACCTGCTATACACTTCTTAAGAAGACTATTAACAGTT	4098
Qy		GlnGlnLysGlnMetProPheCySAspLysAspArgThrLeuMetThrProValGlnAsn	1408
Db		410 CAGCAAAAAACAGATGCTCTTCTGTGACAAAGACAGACATGATATACACTGTGTCAAGAT	4698
Qy		LeuGlnGlnSerTyrMetProProAspSerSerSerProGluAsnLysSerLeuGluVal	1608
Db		470 CTTGACAAAGCCTATGATGCCACTGCACTCTCTCAAGTCCAGAAAAATTAAGTTTGAAGTT	5228
Qy		SerAspThrArgPheHisSerPheSerPheTyrGluLeuLysAsnValThrAsnAspHe	1808
Db		530 AGTCATACAGCTTTTTCACAGTTTTCATTTATGAAATGAAGAATGCACAATTAACCTT	5898
Qy		AspGluArgProIleSerValGlyIleAsnLysMetGlyGluGlyLysPheGlyValVal	2008
Db		590 GATCAACAGACCCATTTCTGTGTGTGTAATTAATAATGGGAGGGAGATTTGGAGTTGTA	6498
Qy		Tyr-LysGlyTyr-ValAsnAsnThrThrValaVala-LysLysLeuValaIleMetVala	2208
Db		650 TATTAAGGGCTAGCTAATATTAACCACTGTGGCAGTGAAGAAAGCTTCAGCAATGGTTG	7098
Qy		SpiLeuThrArgLys-GluLeuLys-GlnGlnPheAsp-GlnGlu-IleLysValMetAla	2388
Db		710 ACATTACTTACTGAAGAACTGAAGAACGACGAGTTTGATCCAAATAAATAAGTATGGCC	7698
Qy		LysCySLeuHis-----GluAsnLeuVal---GluLeuLeuGlyPheSerSer---Asp	2558
Db		770 AAAAGTGTCACACCATGAGAAAAAATTAGTAGGAACCTACTTGGGTTTTCACAGTGAT	8298
Qy		255 Gly 255	
Db		830 GGG 832	
RESULT 2			
US-08-587-889-1			
; Sequence 1, Application US/08587889			
; Patent No. 5654397			
; GENERAL INFORMATION:			
; APPLICANT: CAO, Zhaoan			
; APPLICANT: CROSTON, Glenn E.			
; APPLICANT: GORDELL, David V.			
; TITLE OF INVENTION: INTERLEUKIN-1 RECEPTOR-ASSOCIATED			
; TITLE OF INVENTION: PROTEIN KINASE AND BINDING ASSAY			
; NUMBER OF SEQUENCES: 2			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT			
; STREET: 4 Embarcadero Center, Suite 3400			
; CITY: San Francisco			
; STATE: California			
; COUNTRY: USA			

```

ZIP:4111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/587,889
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-60916
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3590 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-587-889-1

Alignment Scores:
Pred. No.:      6,85e-50      Length:      3590
Score:          515.50       Matches:     156
Percent Similarity: 46.14%   Conservative: 89
Best Local Similarity: 29.38% Mismatches:    192
Query Match:     21.61%      Indels:      94
DB:              1           Gaps:        19

US-10-001-254-16 ((-460) x US-08-587-889-1 (-3590))

QY      4 ProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleAlaGlyLysLeuSer 23
Db      115 CCCCAGCGCCACGACTCTTGTATAGAGAGTCGCCCGCTGGGTCAATGGCCGCTTTACAAG 174
QY      24 AspPheIleAAsPrGlnGlnGlyTrpYalYsLeuAlaValAlaIleLysProSer 43
Db      175 AGT-----GAAGAAGCGCCCTTGAGACCGCGGACTG--GTGCAGTCCCGCCCT-- 222
QY      44 GlyAspAspArg-TyrAsnGlnPheIleArgArgPheGlnAlaLeuLeuGlnThrGl 63
Db      223 -----GATGTGTCGCGAACCAACCGAGCTCGG-----CTGGCGAGCGCTC 264
QY      63 YLysSerProThrSerGlyLeuLeuPheAspTrpGlyThrThrAsnCysThrValGlyAs 83
Db      265 CGGCGACCGCACGCGCGACGCGCTGTGTGCGCTGTGATCAACCGCAACCGCCGCTGTGGCGA 324
QY      83 PleuValAspLeuLeu-----IleGlnAsnGluPh 93
Db      325 CCTGTGTCACATCTTCACGACCTTCAGCTCTCCGTGCGCGGACATCATCAGAGCTTG 384
QY      93 ePheAlaProAlaSerLeuLeuLeuProAspAlaValProLysThrAlaAsnThrLeuPr 113
Db      385 GCACCTTCGCGCCCGCTTCGTCGTCCACAGGACCACTGCCCGGAGCGCCACAGCATCC 444
QY      113 oSer---LysGlnAlaIleThrValGlnGlnLysGlnMetProPheCysAspLysAspar 132
Db      445 TGACACCGCGGAGCGCGAGCGCTGAGAGCCCGGAAATGCCA-----TCTCAGGCTC 498
QY      132 gHrLeuMetThrPro----- 137
Db      499 CACCTTCCTCTCCGCACTTTTCAGAGCTCCAGACCATTACAGAGCTTAGCTCGGCT 558
QY      138 -ValGlnAsnLeuGlnGlnInsertyMeMetProProAspSerserProGlnAsnLysse 157
Db      559 GGTTCCAAGCGCTGCTCCCTGTGGCTCACCACCGCAATTCACAGCCCTTCTTTACCAA 618

```

```

QY 157 r-----leuGlulValSerAspThrArgPheHisSerPh 168
DB 619 GCCAGGCCAGAGAGCTCAGTGTCCCTCTGCAGGAGCCGCCCTCCCTGTTTGCTG 678
QY 168 eSerPheTyrgluleuLeuValThrSerAsnPheAspGluArgProIleSerValgl 188
DB 679 GCCCTCTGTGTAGATTCCCGGGGACCCACAACTTCTGGAGGAG-----724
QY 188 ygluLeuAsnMetGlygluGlyPhegluValValTyrluSgLyTyrrValAsnAsn 208
DB 725 -----CTCAAGATCGGGAGGGTGGCTTTGGTGCCGTACCGGGCGGTATGAGAAC 780
QY 208 rThrValAlaValLysLysLeuAlaAlaMetValAspIleThrThrGluGluLeu 228
DB 781 GGTGTATGCTGTGAAGAGCTGAAGAGAACCGCTGACCTGAGTGAAGTGAAGCA 840
QY 228 nGlnPheAspGlnGluLysValMetAlaLysCysGlnHisGluAsnLeuValGlu 248
DB 841 GAGCTTCTGACGAGGTGAGAGAGCTGTCCAGGTTTCTCCACCCAAACATTGTGACT 900
QY 248 uLeuGlyPheSerSerAspGlyAspAspLeuCysLeuValTyrrValTyrrMet 268
DB 901 TGTGTGCTACTGTGTCTCAGAACGGCTTCTACTGCTGTGTGTGTGTGTGTGTGT 960
QY 268 ySerLeuLeuAspArgLeuSerCys---LeuAspGlyThrProProLeuSerTrpHis 287
DB 961 CTCCCTGAGAGAGACCGTCTCAGTCCAGACCCAGGCTCCACCTCTCTCTGCGCTCA 1020
QY 287 rArgCysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 307
DB 1021 GCGACTGACATCTTGTGGTACAGCCCGGCAATTCACTTCACTCAGAGACAGCC 1080
QY 307 s-----lleHsArgAspLysLysLysLysLysLysLysLysLysLysLys 325
DB 1081 CAGCTCATCATCAGAGACATCAGAGATCAGAGATCAGAGATCAGAGATCAGAG 1140
QY 325 aLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 340
DB 1141 CAGCTGAGAGACTTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 1197
QY 341 -GlnThrValMetThrSerArg-----lleValGlyThrThrAlaTyrrMetAla 357
DB 1198 CCAGAGCAGACATGTGTGGCCGACACAGACAGTGTGGCCGACCTGTGGCTGT 1257
QY 357 oGluAlaLeuArg---GlyGluLysLysLysLysLysLysLysLysLysLys 376
DB 1258 GAGATACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
QY 376 lLeuLeuGluLysLysLysLysLysLysLysLysLysLysLysLysLysLys 395
DB 1318 AGTGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
QY 395 uLeuLeuAspLysLysLysLysLysLysLysLysLysLysLysLysLysLys 407
DB 1378 TCTGAAGAGCTGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
QY 407 uLysThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 411
DB 1438 GAGCAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1497
QY 412 -----AspTyrrLysLysLysLysLysLysLysLysLysLysLysLys 429
DB 1498 CTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1557
QY 429 rTyrrSerValAlaSerGlnCysLeuHisGluLysLysLysLysLysLysLys 449
DB 1558 GGGCCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1617
QY 449 sValGlnGlnLeuLeuGlnLysMetThrAla 459
DB 1618 GGTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1648

```

RESULT 3

```

US-09-016-434-1093
; Sequence 1093, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERBWITH
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1093:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1220312
; US-09-016-434-1093
; Alignment Scores:
; Pred. No.: 6,85e-50 Length: 3590
; Score: 515.50 Matches: 156
; Percent Similarity: 46.14% Conserved: 89
; Best Local Similarity: 29.38% Mismatches: 192
; Query Match: 21.61% Indels: 94
; DB: 4 Gaps: 19
; US-10-001-254-16 (1-460) x US-09-016-434-1093 (1-3590)
QY 4 ProIleThrProSerThrTyrrValArgCysLeuAsnValGlyLeuLysLeuSer 23
DB 115 CCCGCGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 174
QY 24 AspPheIleAspProGlnGlnGlyTyrrLysLysLysLysLysLysLysLys 43
DB 175 AGT-----GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 222
QY 44 GlyAspAspArg-TyrrAsnGlnPheHisLysArgArgPheGluAlaLeuGln 63
DB 223 -----GATCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 264
QY 63 yLysSerProThrSerGlnLeuLeuLeuPheAspTrpGlyThrThrAsnCythrVal 83
DB 265 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 324

```

[illegible][illegible]

DB: 5 Gaps: 19
US-10-001-254-16 (1-460) x PCT-US96-09193-1 (1-3590)
QY 4 ProIlethrProserThrTyValArgCysLeuAnValGlyLeuIleArgIysLeuSer 23
DB 115 CCCCCGGCCGAGCACTTCTGTACAGAGTCCGCCCTGGGATCATGTGCGCTTTACAA 174
QY 24 AspPheIleAspProGlnGluIlyTyPryblybLeuAlaValAlaIleIlyblybProser 43
DB 175 AGT-----GATGACAGCCCTGGAGCCCGCCGACTG---GTCCAGTTGCCGCCCT--- 222
QY 44 GlyAspArgArg-TyValGlnPheHisIleArgArgPheGluAlaLeuGlnIleThrG1 63
DB 223 -----GATCGTGGCCGACACAGCCGAGCTGGG-----CTGTGGACAGCGCTC 264
QY 63 YLysSerProThrSerGlnLeuLeuPheAspPheIlyThrThrAspCysThrValGlyAs 83
DB 265 CCGGACGCGCACGGGACAGCTCTGTGGCTGTGATCAACCGCAAGCCCGCTGTGGCCGA 324
QY 83 PLeuValAspLeuLeu-----IleGlnAsnGluPb 93
DB 325 CCGTGTGACATCTCTACAGCACTGACGTCTCGTGGCGGACATCATACAGCCTG 384
QY 93 ePheAlaProAlaSerLeuLeuLeuProAspAlaValProlybThrAlaAsnThrLeuPr 113
DB 385 GCACCCCTCCCGCCGCTTCCGTCCCGACAGCACACTGCCCGGAGCCCGACAGCATCC 444
QY 113 oSer---LybGlnAlaIleThrValGlnGlnIlybGlnMetProPheCysAspIlybAspAr 132
DB 445 TGACACCCGCGACCGGAGGCTGGAGCCCGGAGAGTTGCCA-----TCCTGACCTTC 498
QY 132 gThrIleuMetThrPro-----LeuGluValSerAspThrArgPheHisSerPh 168
DB 499 CACCTTCCTCTCCAGCTTTTCCAGGCTCCCAAGCCCATTCAGGCGCTGAGCTCGGCT 558
QY 138 -ValGlnAsnLeuGlnIlybSerTyMetProProAspSerSerProGluAsnIlybSe 157
DB 559 GGTTCACAGCCCTGCTCTGTGGCTTCACCGGCATCTCCAGCCCTTCTTACCA 618
QY 157 r-----LeuGluValSerAspThrArgPheHisSerPh 168
DB 619 GCCAGGCCGACAGAGCTGATGCTCCCTCTGACGGAGACCCGCCCTTCGATTGTGCTG 678
QY 168 eSerPheTyGlnLeuIlybAsnValThrAsnAspPheAspGluArgProIleSerValG1 188
DB 679 GCCCTCTGTAGATTTCCGGGGGACCCACACATCTTCGAGAGAG----- 724
QY 188 yGlyAsnIlybMetGlyGlnGlyGlyPheGlyValIlybTyTybGlyTyTyValAsnAsnTh 208
DB 725 ----CTCAAGATCGGGAGGGGTGGTTGGGTGGGTGATACCGGGCGGTGATAGAGAAC 780
QY 208 rThValAlaValIlybIlybLeuAlaAlaMetValAspIleThrThrGlnGluLeuIlybG1 228
DB 781 GGTGTATGCTGTGAAGAGCTGAAGAGAGAACCTGACCTGAGTGAAGTGAAGCA 840
QY 228 nGlnPheAspGlnGlnIlybIlybValMetAlaIlybCysGlnIlybAsnLeuIlybG1 248
DB 841 GAGCTTCCTGACCGAGGTGAGAGCACTGTCCAGGTTTGTACCCAAACATTTGGAGATT 900
QY 248 uLeuGlyPheSerSerAspGlyAspAspLeuCysLeuValTyTyValTyMetProAsnG1 268
DB 901 TGCTGCTACTGTCTCAGAACGGCTTCTACTGCTGTGTATACGGCTTCGCGCCAAAG 960
QY 268 ySerLeuLeuAspArgLeuSerCys---LeuAspGlyTyThrProPoleuSerTyPheIlybSe 287
DB 961 CTCCTGTGAGGACCGTCTCCACTGACAGCCAGGCTTCGACACTCTCTCTGCGCTCA 1020
QY 287 tArgCysIlybIleAlaGlnGlyAlaAlaAsnGlyIlyLeasnPheIlybGlnAsnHisI 307
DB 1021 GCGACTGACATCTCTTGGGTATACAGCCCGGCAATTCAGTTTCTACATTCAGAGCC 1080
QY 307 s-----IleHisArgAspIlybSerSerAlaAsnIleLeuLeuAspGluAlaPheThrAl 325

DB 1081 CAGCTCATCCATGAGACATCAAGAGTTCCAAAGCTCTTGTGATGAGAGCTGACACC 1140
QY 325 alybIleSerAspPheGlyLeuAlaArgAlaSerGlnIlybPheAla----- 340
DB 1141 CAGCTGGAGACTTTGGCTTGGCCCGGTTGAGC---CGTTTGGCGGTCCAGCCCGAG 1197
QY 341 -GlnThrValMetThrSerArg-----IleValGlyThrThrAlaTyMetAlaPr 357
DB 1198 CCAAGACAGCATGTGTGGCGCCGACACAGACAGTCCGGGACCTCGGCTTACCTGCCGA 1257
QY 357 oGluAlaLeuArg---GlyGlnIleThrProlybSerAspIleTySerPheGlyValIy 376
DB 1258 GAGATCATCAAGACGGAGAGCTGTGTCGACAGGACACCTTCAGCTTTGGGGTGT 1317
QY 376 lLeuGlnGlnIleleThrGlyLeuProAlaValAspGlnHis---ArgGluProGlnI 395
DB 1318 AGCTGTAGAGCTTGTGTGTGACAGAGGCTGTGAAGACGACCGGTGCCAGACCAAGTA 1377
QY 395 uLeuLeuAspIleIlybGlnGlnIlybGluAspGlu-----G1 407
DB 1378 TCTGAAGACCTGTGTGAAGAGAGGCTGAGAGAGCTGTGAGTGTGAGAGACACCA 1437
QY 407 uLybThrIleGlu----- 411
DB 1438 GAGACACTGCACAGCAGGTCTGCTGACAGATCCTGGGCTGTCCATCGCATGACAGAT 1497
QY 412 -----AspTyTrIleAspIlybIlybMetAsnAspAlaAspSerThrSerValGluAlaMe 429
DB 1498 CTACAGAAAGACCTGACCCAGCCCGGCGCTGCCCATTCAGCTGGCGCTGGGCT 1557
QY 429 tTySerValAlaSerGlnCysLeuHisGluIlybAsnIlybAspProAspIleIlybIy 449
DB 1558 GGGCAGCTGGCTGTGCTGTCTGTCTGACCGCCGGCCAAAAGAGGCTCTTATGACCA 1617
QY 449 sValGlnGlnLeuLeuGlnIlybMetThrAla 459
DB 1618 GGTGTACGAGAGCTTGAAGAAAGTGCAGGCA 1648

RESULT 5
US-09-579-182-4
; Sequence 4, Application US/09579182
; Patent NO. 650628
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USBS THEREFOR
; FILE REFERENCE: NMI-161
; CURRENT APPLICATION NUMBER: US/09/579,182
; CURRENT FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1488
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-579-182-4

Alignment Scores:
Pred. No.: 6,19e-46 Length: 1488
Score: 476.50 Matches: 130
Percent Similarity: 54.01% Conservative: 52
Best local Similarity: 38.58% Mismatches: 104
Query Match: 19.97% Indels: 51
DB: 4 Gaps: 14
US-10-001-254-16 (1-460) x US-09-579-182-4 (1-1488)
QY 145 TyMetProProAspSerSer-----SerProGluAsnIlybSerLeuGluVal 160
DB 55 TCTCTTAATTCGATTAATCAAGATTCGACGAGGTCTGTGAAGAGAAAG---GAGCTA 111
QY 161 SerAspThrArg-----PheHisSerPheSerPheTyGlu 172

```

Db      112 ACTGCTCCAAAAGAGGCGCTACTGCGCATATTGCTGCACAAACCTTACTTCCGAGAG 171
Qy      173 LeuLysAsnValThrAsnAspPheAspGluArgProIleSerValGlyGlyAsnLysMec 192
Db      172 TTAGCTGCCCACTAAAACCTT-----CGAGCAAGATGCTT-----CTT 213
Qy      193 GlyGluGlyGlyPheGlyValValTyrLysGlyTyrValAsnAsnThr-----ThrVal 210
Db      214 GGAAGAAGAGGTTTCCGACCGTTTACAAAGTGTGTAGAGACCAACAGACAGATAGTA 273
Qy      211 AlaValLysLysLeuAlaAlaMetValAspLleThrThrGluGluLeuLysGlnPhe 230
Db      274 GCGTTAAACAGCTT-----GATCGAAACGCTGTACAGGAAACAGAGAGCTTT 321
Qy      231 AspGlnGluLysValMetAlaLysCysGlnHisGlnAsnValGluLeuGly 250
Db      322 CTGTAGAGGTTCTTATGCTGAGCCTTCTGCATCATCCCATCTTGGAATTGATTTGCT 381
Qy      251 PheSerSerAspGlyAspAspLeuValLeuValTyrMetProAsnGlySerLeu 270
Db      382 TATTGTGCTATGGGAGCCAGCGCTTCTGTGTATGATATAGTATAGCACTAGATCATTTG 441
Qy      271 LeuAspArgLeuSerCysLeu---AspGlyThrProProLeuSerTyrHisMetArgCys 289
Db      442 GAGGATCATCTACACGATCTTCCACCTGATTAAGAGCCTCTAGACTGAGATCAATG 501
Qy      290 LysLleAlaGlnGlyAlaAlaAsnGlyLleAsnPheLeuIleGluAsnHis----- 307
Db      502 ACAATACCGGACAGAGACGAAAGGAGCTGAGATGTGATGTGAATGAAGCAACGACCT 561
Qy      308 ---LleHisArgAspLysSerAlaAsnLleLeuAspGluAlaPheThrAlaLys 326
Db      562 GTGATCTACAGAGCCTGAATATCTCAACATTTCTTCGATAGGCTATCACCCAAAG 621
Qy      327 IleSerAspPheGlyLeuAlaArgAlaSerGluLysPheAlaGlnThrValMetThrSer 346
Db      622 TTATCTGATTTTGGGTAGCTAGTATAGTCCCGTGGCGATAAACACATGTCTCACT 681
Qy      347 ArgLleValGlyThrThrAlaTyrMetAlaProGlu---AlaLeuAspGlyGluLleThr 365
Db      682 CGGTGATGAGGACATATGTGTTATGTGCACCGGAATATGTCAGACAGACGCAACTACA 741
Qy      366 ProLysSerAspLysIleTyrSerPheGlyValValLeuLeuGluLleIleThrGlyLeuPro 385
Db      742 TTGAAATCCGATGTTTATAGCTTTGGGGTGTGTCTTCGAGCTCATACAGGGTCAAAA 801
Qy      386 AlaValAspGluHisArgGluPro-----GlnLeuLeuAspLleLysGluGlu 402
Db      802 GCTATTGATATATGCTCGAGCACCCGAGACACAAACCTTGTCCGATGGGCTAGGCCCTTG 861
Qy      403 IleGluAspGluGlySerThrIleGluAspTyrLleAspLysLysMetAsnAspAlaAsp 422
Db      862 TTCAAATGCTAGAAAAGTTT-----CCGAAGATGGCGGAT----- 897
Qy      423 SerThrSerValGluAlaMetTyr-----SerValAla 433
Db      898 ---CGATCGCTGCAAGGCGGTATCCATGCGTGTCTATATCAAGACATTCGAGTTGCA 954
Qy      434 SerGlnCysLysHisGlyLysValAsnLysAlaArgProAspLleLysVal 450
Db      955 GCATATGTTTACAGAAACAGCAAGCAACCAAGACCACTGATTTGCGACGTTG 1005

```

RESULT 6

```

US-09-228-986-10
; Sequence 10, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020

```

```

; CURRENT APPLICATION NUMBER: US/09/228, 986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 2336
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-228-986-10

Alignment Scores:
Pred. No.:      8.44e-43      Length:      2336
Score:          452.50      Matches:      124
Percent Similarity: 51.93%      Conservative: 51
Best Local Similarity: 36.80%      Mismatches: 115
Query Match:    18.96%      Indels:      47
DB:              4          Gaps:      13

US-10-001-254-16 (1-460) x US-09-228-986-10 (1-2336)
Qy      132 ArgThrLeu---MetThrProValGlnAsnLeuGlnSerTyrMetProAspSer 150
Db      967 AGAACACTTAACACGGAACGCTCGAAGATGGAGACAGAGATCTGGCC----- 1017
Qy      151 SerSerProGluAsnLysSerLeuGluValSerAspThrArgPheHisSerPheSerPhe 170
Db      1018 -----CATAGATTGATTTAC 1032
Qy      171 TyrGluLeuLysAsnValThrAsnAspPheAspGluArgProIleSerValGlyGlyAsn 190
Db      1033 AAGAGCTCGGTATTCGCCACGAGGAGGTTTCGGAGCAAA-----AAC 1074
Qy      191 LysMetGlyGlyGlyPheGlyValValTyrLysGlyTyrVal-----AsnAsnThr 208
Db      1075 CTTTGGGGTACGGGGGATTTGGCATGTTTACAAAGGTTTCTCCCGAGAGCGGCCAA 1134
Qy      209 ThrValAlaValLysLysLeuAlaAlaMetValAspLleThrGlu-----GluLeu 226
Db      1135 GAAGTCGACAGTGA---TGATTAACGACGAGGATTCAGAGAAAGCA 1176
Qy      227 LysGlnGlnPheAspGlnGluLysValMetAlaLysCysGlnHisGlnAsnLeuVal 246
Db      1177 ATTAAGGGGTTTTCGACAGATCTCAAGCATGGGCGCTACAGACCGGAACCTGGTT 1236
Qy      247 GluLeuGluLysPheSerSerAspGlyAspAspLeuValTyrValTyrMetPro 266
Db      1237 CAACCTCGAGAGATGGTCCGAGGACATACACAGCTTATCATCTTTACAGACTACATGCC 1296
Qy      267 AsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThrProProLeuSerTyrHis 286
Db      1297 AACGGAAGC---CTGCATTAACATCATCTTCGGTAGTCCGACACAGTCCGCGTGGCAT 1353
Qy      287 MetArgCysLysLleAlaGlnGlyAlaAlaAsnGlyLleAsnPheLeuHisGlnAsnHis 306
Db      1354 CGCGCATACCGCATCTTAAGGGCGTAGCACCGGGCGTGTGTATGTGCACGAGCAATGG 1413
Qy      307 His-----LleHisArgAspLysSerAlaAsnLleLeuLeuAspGluAlaPhe 323
Db      1414 GAGAAGAGGTCGTCACACAGGACATTAAGTTCAGACACATGCTGTGGATTGGAGCTTC 1473
Qy      324 ThrAlaLysLleSerAspPheGlyLeuAlaArgAlaSerGluLysPheAlaGlnThrVal 343
Db      1474 AACGGGCGCTTAGTACTTTCGGGCTCGCTCGGCTGTATGAT---CACAGTGAATCCG 1530
Qy      344 MetThrSerArgLleValGlyThrThrAlaTyrMetAlaProGluAlaLeuArg---Gly 362
Db      1531 GAGACGACATATGTGTGAGAACGTTGGGGTATCATGACACCGGAGTTGATACAAACGGGG 1590
Qy      363 GluLleThrProLysSerAspLysIleTyrSerPheGlyValValLeuLeuGluLleIleThr 382
Db      1591 AAGGCACTCTTACGTCGAGCGTTCAGTTCGGTGTGCTGTGGAGGGTGGCTTGC 1650
Qy      383 GlyLeuProAlaValAsp-----GluHisArgGluProGlnLeuLeuAspLleLys 400

```


US-09-228-986-7
 ; Sequence 7, Application US/09228986
 ; Patent No. 6359198
 ; GENERAL INFORMATION:
 ; APPLICANT: Strabala, Timothy
 ; APPLICANT: Niemenhuizen, Niels
 ; TITLE OF INVENTION: Compositions Isolated from Plant Cells
 ; FILE REFERENCE: 11000/1020
 ; CURRENT APPLICATION NUMBER: US/09/228,986
 ; CURRENT FILING DATE: 1998-01-12
 ; NUMBER OF SEQ ID NOS: 130
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 7
 ; LENGTH: 2432
 ; TYPE: DNA
 ; ORGANISM: Pinus radiata
 US-09-228-986-7

Alignment Scores:
 Pred. No.: 2,656-42 Length: 2432
 Score: 448.50 Matches: 113
 Percent Similarity: 56.83% Conservative: 66
 Best Local Similarity: 35.87% Mismatches: 109
 Query Match: 18.80% Indels: 27
 Gaps: 11

US-10-001-254-16 (1-460) x US-09-228-986-7 (1-2432)
 QY 154 GUAASLysSerLeuGluValSerAspThrArgPheHisSerPheSerPheTyrGluLeu 173
 DB 1028 GAACAGACACATGTTGATTTCTAGGCGCATTTGAAGATCTCATTCMAAGAGTTA 1087
 QY 174 LysAsnValThrAsnAspPheAspGluArgProIleSerValGlyAsnLeuMetGly 193
 DB 1088 CGGTTCACATTAATTTAGTACAG-----AATATTTAGA 1129
 QY 194 GUGLYGlyPheGlyValValTyrGlyGlyTyrVal---AsnAsnThrThrValAlaVal 212
 DB 1130 GTAGAGATATGATGATTTGCTATTAAGATCTCTACAAAGTGCACATATGATGACATA 1189
 QY 213 LysLeuLeuAlaAlaMetValAspIleThrThrGluGluLeuLysGlnInPheAspGln 232
 DB 1190 AAAAGTTTGA-----GATGTAATGTCGAGAGAGAAATTCATTTCAACA 1240
 QY 233 GluIleLysValMetAlaLysCysGlnHisGluAsnLeuValGluLeuGlyPheSer 252
 DB 1241 GAGGTGAAATGATCAGCTTGGCTGTCATAGAACCTATTACATTTGATGATTTTGC 1300
 QY 253 SerAspGlyAspAspLeuCysLeuValTyrValTyrMetProAsnGlySerLeuLeuAsp 272
 DB 1301 ACAACCTCCAGAGAGAGGCTTGTGTATCCCTACACATGCAATGGAAGTGTG----- 1354
 QY 273 ArgLeuSerCysLeu-----AspGlyThrProPheLeuSerTyrPheMetArg 288
 DB 1355 ---GCCCTTGTCTTACAGATCATATTATTAAGAAAGCTTGGCTGAGCTGCTTACTGCG 1411
 QY 289 CysLysIleAlaGlnGlyAlaAlaAsnGlyLysLeuPheLeuHisGluAsn----- 305
 DB 1412 AAGGTAAGCTCTTGGAGCGCTGAGGAGCTGTAATTTTGCATGAGCAATGTGATCCC 1471
 QY 306 HisHisIleHisAspIleLysSerAlaAsnIleLeuLeuAspGluAlaPheThrAla 325
 DB 1472 AAGATTATTCACCGGATGTGAAGACAGCAATATATTACTGATGTAATTTTGAAGCT 1511
 QY 326 LysIleSerAspPheGlyLeuAlaArgAlaSerGluLysPheAlaGlnThrValMetThr 345
 DB 1532 GTTGTTCAGATTTTGGGTTCGCAAGCTCTTGAT---CACAGGATTCCTCATGTGACT 1588
 QY 346 SerArgIleValGlyThrThrAlaTyrMetAlaProGluAlaLeu---ArgGlyGluLe 364
 DB 1589 ACTGCTTTCGAGGAGCGTAGTGCATTCCTCCAGAAATACCTTCAACGGGACCAATCT 1648

QY 365 ThrProLysSerAspIleTyrSerPheGlyValValLeuLeuGluIleIleThrGlyLeu 384
 DB 1649 TCAGAGAAACATGATATTTGGCTTGGATATTAAGTATGATGACATTCACAGACAA 1708
 QY 385 ProAlaValAsp-----GluHisArgGluProGlnLeuLeuLeuAspIleLysGlu 401
 DB 1709 AGGCTTTAGATTTTGGCCAGGCTGCAGAACCAAAAGTTATAGCTGATGGGTGAAAG 1768
 QY 402 GluIleGluAspGluLysThrIleGluAspTyrIleAspLysMet---AsnAsp 420
 DB 1769 AAGCTTCAT---CAAGAGAAAGTTGCACCTCTGCTGATTAAGATCTTAAGGGCAAT 1825
 QY 421 AlaAspSerThrSerValGluAlaMetTyrSerValAlaSerGlnCysLeuHisGluLys 440
 DB 1826 TTGTATAGAGTTGAGTAGAAGAGATGATGCTTTCAGGTTTCTGTACAGCCCAATTCAG 1885
 QY 441 LysAsnLysAspProAspIleLysLysValGlnGluLeuGln 455
 DB 1886 CTTGACATGCTCCAAAATGCTGATGTTTGAGAAATGTGGA 1930

RESULT 9
 US-09-135-232-1
 ; Sequence 1, Application US/09135232
 ; Patent No. 6262228
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Zhaoan
 ; TITLE OF INVENTION: IRAK3 Polypeptides, Polynucleotides and Methods
 ; FILE REFERENCE: 798-019
 ; CURRENT APPLICATION NUMBER: US/09/135,232
 ; CURRENT FILING DATE: 1998-08-17
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 2288
 ; TYPE: DNA
 ; ORGANISM: human
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (64)..(1851)
 US-09-135-232-1

Alignment Scores:
 Pred. No.: 4,1e-41 Length: 2288
 Score: 438.00 Matches: 127
 Percent Similarity: 44.80% Conservative: 84
 Best Local Similarity: 26.96% Mismatches: 188
 Query Match: 18.36% Indels: 72
 Gaps: 12

US-10-001-254-16 (1-460) x US-09-135-232-1 (1-2288)
 QY 16 ValGlyLeuIleLysLysSerAspPheIleAspProGlnGluGlyTyrLysLysLeu 35
 DB 136 CTGAGAGCTCTGCGCTGTTCTGACAGCTGCGAGCGCGCTGCGCTGCGCGCTG 195
 QY 36 AlaValAlaIleLysLysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgArg 55
 DB 196 GCAGAGAGATTTCAAGC-----AGCTGGCGATGATTCGTCAT 234
 QY 56 PheGluAlaLeuGlnThrGlyLysSerProThrSerGluLeuLeuPheAspTyrGly 75
 DB 235 ATTGAAGAATGATTAACCAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAG 294
 QY 76 ThrThrAsnCysThrValGlyAspLeuValAspLeu----- 87
 DB 295 CAGAAAAACAAAGCCATCGGTGACCTTTTACAGGCTCTCCAGAGATGGGACATCTGCA 354
 QY 88 -----LeuIleGlnGlnGluPhePheAlaProAlaSerLeuLeuLeuProAspAla 104
 DB 355 GCATTCATTTATTTCAAACTAT-----GAGCAATGTTGATGCTTCAGAG 402
 QY 105 ValProLysThrAlaAsnThrLeuProSer-----LysGluAla-----IleThr 119

```

Db 403 AAGATTATCAGGAAGTCGATTCCAAATATATTATTCACGAAACAGCCAAATGTCACC 462
Qy 120 ValGlnGlnIlysgIleMetProPheCyAspIlyAspArgThrLeuMetThrProValGln 139
Db 463 GTGGATTAATGTTCTTATTCCTGACATATATGAAAAAGAGTACTGCTT----- 510
Qy 140 AsnLeuGlnIleSerThrMetProAspSerSerSerProGluAsnIlySerLeuGln 159
Db 511 -----AAATCT----- 516
Qy 160 ValSerAspThrArgPheHisSerPheSerPheThrGluLeuIlyAsnValThrAsn 179
Db 517 -----TCCATCAGCTTCCAAATATATGAAAGAACTTGAAAT 555
Qy 180 PheAspGluArgProIleSerValGlyIlyAsnIlyMetGlyGlyGlyPheGlyVal 199
Db 556 TTCACAAAGACTTCTTA-----ATTGGAGAGAGAGAGATTTTGGAG 597
Qy 200 ValIlyIlysgIlyThrValAsnAsnThrThrValAlaValIlyIlyLeuAlaIleMetVal 219
Db 598 GTATACAGAGTCGAGATTCAAACCTAACATATGCTGCAATTTTAAACAGAGAGAA 657
Qy 220 AspIleThrThrGluGluLeuIlysgIlnIlePheAspGlnIlyIleValMetAlaIly 239
Db 658 AAAATGCAGTGTAAAGACATTTGAAAGGTTTATCTGAGCTTGAAAGTTTACTACTG 717
Qy 240 CysGlnIlysgIlyAsnLeuValGluLeuLeuGlyPheSerSerAspGlyAspAspLeu 259
Db 718 TTTTATTCACCAACATCTAGAGTTGGCTGCTATTTTACAGAGACTGAGAACTTCTGT 777
Qy 260 LeuValIlyValIlyMetProAsnIlySerLeuLeuAspArgLeuSerCysLeuAspGly 279
Db 778 CTGATTATTCATACATAGTGAATGGAACACTTTTGAACAGTTGACAGTGTCTAGTAC 837
Qy 280 ThrProPheLeuSerThrPheMetArgCysIlyIleAlaGlnIlyAlaIleAsnIly 299
Db 838 ACGGCCCACTCCCTTGGCACATTCGAATCGGTATATTAAATGAAATTCACAAAGCCATT 897
Qy 300 AsnPheLeuHisGluAsnHis-----IleHisArgAspIleIlySerIleAsn 316
Db 898 CACTACCTGCACACAGTTCCAAACCATGCTCGGCTGCTGGCAGTATTCAGAGTCCAAAC 957
Qy 317 IleLeuLeuAspGluAlaPheThrAlaIlyIleSerAspPheGlyLeuAlaArgIleSer 336
Db 958 ATTCCTTTGGATGATCACTTTCACACCAACCTAATGATTTTGCATGCGACACTCCGG 1017
Qy 337 GluIlyPheAlaGlnThrValMetThrSerArgIleValGlyThrAla----- 353
Db 1018 TCCCACTAGAACATCAGAGTTGTACATTAATATGACACAGCAGCAGATTAACATCTG 1077
Qy 354 ---TyrMetAlaProGluAlaLeuArg---GlyGluIleThrProIlySerAspIleTyr 371
Db 1078 TGGTACATGCGCAGAGATACATCAGACAGGAGAACTTTCATTAAACAGATGTCTAC 1137
Qy 372 SerPheGlyValIleLeuLeuGluIleIleThrGlyLeuProAlaValAspGluHisArg 391
Db 1138 AGCTTTGGAATTTGTAATATGAAAGTTCTTAACAGATGTAGAGTATGATGATCCA 1197
Qy 392 GluProGlnLeuLeuLeuAspIleIlysgIlnIleGluAspGlnIly-----LysThr 409
Db 1198 AAACATATCCAGCTCGGAGTCTCTTAGAGATGATGAGAGAGAGCGCTGGATTC 1257
Qy 410 IleGluAspIlyIleAspIlyIlyMetAsnAspAlaAspSerThrSerValGluAlaMet 429
Db 1258 TGTCTTCATTTCTAGATTAAGAAAGTCTCTCCCTGCGAATTTCTGCGCAAGCTC 1317
Qy 430 TyrSerValAlaSerGlnCysLeuHisGluIlyIlyIlyAsnIlyAspProAspIleIly 449
Db 1318 TTCTGTTTGGAGCGCGGTGTGTCGCAACGCGGCAAAAGTTAAAGCACTCAATGATGAA 1377
Qy 450 ValGlnGlnLeuLeuGlnIleMetThrAlaSer 460
Db 1378 GTTTTAAATACTCTTGAAGTACTCAAGCCAGC 1410

```

```

RESULT 10
US-09-863-549-1
; Sequence 1, Application US/09863549
; Patent No. 6576444
; GENERAL INFORMATION:
; APPLICANT: Cao, Zhaoan
; TITLE OF INVENTION: IRAK3 Polypeptides, Polynucleotides and Methods
; FILE REFERENCE: T98-019
; CURRENT APPLICATION NUMBER: US/09/863,549
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 09/135,232
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2288
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (64)..(1851)
; US-09-863-549-1

Alignment Scores:
Pred. No.: 4,1e-41 Length: 2288
Score: 438.00 Matches: 127
Percent Similarity: 44.80% Conservative: 84
Best Local Similarity: 26.96% Mismatches: 188
Query Match: 18.36% Indels: 72
DB: 4 Gaps: 12

US-10-001-254-16 (1-460) x US-09-863-549-1 (1-2288)
Qy 16 ValIlyLeuIleArgIlyLeuSerAspPheIleAspProGlnGlnIlyIlyIlyLeu 35
Db 136 CTCGAGAGCTCTCGCTGTTCTGACAGCTGCGAGCGCGGCTGGCGGCGCTG 195
Qy 36 AlaValAlaIleIlyIlyIlyProSerGlyAspAspArgIlyIleGlnIlePheHisIleArg 55
Db 196 GCAGAGAGACTTTCACAC-----ACCTGCTGAGATGTCGTCAT 234
Qy 56 PheGlnAlaLeuLeuGlnThrGlyIlySerProThrSerGluLeuLeuPheAspIlyIly 75
Db 235 ATTGAAGAATGATGACCAAGGTAAGAGTGAACAAGAAATTAATCTTGTCTGGCA 294
Qy 76 ThrThrAsnCysThrValGlyAspLeuValAspLeu----- 87
Db 295 CAGAAAAACAAGACCATCGGTGACCTTTACAGGTCTCCAGAGATGGACATCGTGA 354
Qy 88 -----LeuIleGlnAsnGluPhePheAlaProAlaSerLeuLeuProAspAla 104
Db 355 GCTATTCATTTAATTAACTAATCTAT-----GAGACAGAGTTGAGTCTTCAAG 402
Qy 105 ValProIlyThrAlaAsnThrLeuProSer-----LysGluAla-----IleThr 119
Db 403 AAGATTATCAGGAAGTCGATTCCAAATATATTATTCACGAAACAGCCAAATGTCACC 462
Qy 120 ValGlnGlnIlysgIleMetProPheCyAspIlyAspArgThrLeuMetThrProValGln 139
Db 463 GTGGATTAATGTTCTTATTCCTGACATATATGAAAAAGAGTACTGCTT----- 510
Qy 140 AsnLeuGlnIleSerThrMetProAspSerSerSerProGluAsnIlySerLeuGln 159
Db 511 -----AAATCT----- 516
Qy 160 ValSerAspThrArgPheHisSerPheSerPheThrGluLeuIlyAsnValThrAsn 179
Db 517 -----TCCATCAGCTTCCAAATATATGAAAGAACTTGAAAT 555
Qy 180 PheAspGluArgProIleSerValGlyIlyAsnIlyMetGlyGlyGlyPheGlyVal 199
Db 556 TTCACAAAGACTTCTTA-----ATTGGAGAGAGAGATTTTGGAG 597

```

```

Qy 200 ValTyrIysglYrValAsnThrThrValAlaValIysIysLeuAlaAlaMetVal 219
Db 598 GTATACGAGTGGAGATTCAAAACCTTAACATATGCTGTCAAAATTATTTAAACAGAGAAA 657
Qy 220 AspIleThrThrIleGluIleuIysGlnIlePheAspGlnIleIysValMetAlaIys 239
Db 658 AAATGAGTGTAAAGACATTTGGAAGAGGTTTATCTGTGCTTGAAGTTTCTACTG 717
Qy 240 CysGlnHisGluAsnLeuValGluLeuLeuIysPheSerSerAspGlyAspAspLeuCys 259
Db 718 TTTCATACCCCAACATACCTAGAGTTGGCTGCATATTTTACAGAGACTGGAAGCTTCTGT 777
Qy 260 LeuValTyrValTyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGly 279
Db 778 CTGATTATTCATATCATGAGAAATGGAACACTTTTTCAGATTTGCAAGTGTGTAGGAC 837
Qy 280 ThrProLeuSerTrpHisMetArgCysIysIleAlaGlnIleAlaIleAsnGlyIle 299
Db 838 ACGGCCCCACTCCCTTGGCAGATTCGATATATTAATGAAGAAATATCCAAAGCCATT 897
Qy 300 AspPheLeuHisGluAsnHisHis-----IleHisArgAspIleIysSerAlaAsn 316
Db 898 CACTACCTGCAACAAGTTCAACATGCTGCTCATCTGTGCACTATATATCAAGTGAAC 957
Qy 317 IleLeuLeuAspGluAlaPheThrAlaIysIleSerAspPheGlyLeuAlaArgAlaSer 336
Db 958 ATCTTTTGGATGATGATTCATTCACCAACATACTGATTTTGGCATAGGCACACTTCGG 1017
Qy 337 GluIysPheAlaGlnThrValMetThrSerArgIleValGlyIleThrAla----- 353
Db 1018 TCCCACTAGAACATCAGAGTTGTACCAATAATATGACCAACAGACAGTAAACATCTG 1077
Qy 354 ---TyrMetAlaProGluAlaLeuArg---GlyGluIleThrProIysSerAspIleTyr 371
Db 1078 TGGTACATGCCAAGAGATACATACAGACAGGGAACCTTTCATTAACAGATGTCTAC 1137
Qy 372 SerPheGlyValValLeuLeuGluIleIleThrGlyLeuProAlaValAspGluHisArg 391
Db 1138 AGCTTGGAAATGTAAATATGAAAGTTCTAACAGAGTGAAGTGAATGATGATGCCA 1197
Qy 392 GluProGlnLeuLeuLeuAspIleIysGlnIleGluIleGluIleGluIleGluIle 409
Db 1198 AAACATATTCAGCTCGCGGATCTCTTGAAGAAATGTATGGAAGAGAGCGCTGATTC 1257
Qy 410 IleGluAspTyrIleAspIysIysMetAsnAspAlaAspSerThrSerValGluAlaMet 429
Db 1258 TGCTCTCATTTCTAAGTAAGAAAGTGCCTCCCTCGGAATTTCTGCGCAAGCTC 1317
Qy 430 TyrSerValAlaSerGlnCysLeuHisGluIysIysAsnIysArgProAspIleIysIys 449
Db 1318 TTCTGTTTGGAGCGCGGTGTGTGCGCAACGGGGCAAGTAAACATCAATGATGAGA 1377
Qy 450 ValGlnGlnLeuLeuGlnIleMetThrAlaSer 460
Db 1378 GTTTTAATACTCTTGAAGTACTCAAGCCAGC 1410

```

```

; TYPE: DNA
; ORGANISM: Pinus radiata
; US-09-228-986-4

Alignment Scores:
Pred. No.: 3e-40 Length: 2868
Score: 432.00 Matches: 113
Percent Similarity: 55.35% Conservative: 63
Best Local Similarity: 35.53% Mismatches: 112
Query Match: 18.11% Indels: 30
DB: 4 Gaps: 12

US-10-001-254-16 (1-460) x US-09-228-986-4 (1-2868)

Qy 153 ProGluAsnIysSerLeuGluValSerAspThrArgPheHisSerPheSerPheTyrGlu 172
Db 1071 CTGCTGAGAGAGATCCAGAAAGTTCACTTAAGGCACTTAAGAGGTTCTCATTAACGGAA 1130
Qy 173 LeuIysAsnValThrAsnPheAspGluArgProIleSerValGlyIleAsnIysMet 192
Db 1131 TTACAGGTTCACTGATGGTTTACCAATAGA-----AACATTCTT 1172
Qy 193 GlyGluGlyIysPheGlyValValTyrIysGly---TyrValAsnAsnThrThrValAla 211
Db 1173 GGCAGAGGTGTGTTTGGAAAGGTGTACAAAGGCGCTTGCAATGCTTCTGTGGTGGCT 1232
Qy 212 ValIysLeuSerLeuAlaAlaMetValAspIleThrThrGluIleuIysGlnIlePheAsp 231
Db 1233 GTTAAACGTCTGAAG-----GAAGAGCGTACACCGGTGGAGAGATTGCAAGTTTCAA 1283
Qy 232 GlnGluIleIysValMetAlaIysCysGlnHisGluAsnLeuValGluLeuIysPhe 251
Db 1284 ACAGAAAGTGAAGTGAATAGCATGAGCAATATAGGAACCTCTTGACATACGTGAATTC 1343
Qy 252 SerSerAspGlyAspAspLeuCysLeuValTyrValTyrMetProAsnIysSerLeu 271
Db 1344 TGCATGACACCCACTGAACGCGCTGTGTATCCCTCAATGCGCAATGGAAGTGT--- 1400
Qy 272 AspArgLeuSerCysLeu-----AspGlyThrProProLeuSerTrpHisMet 287
Db 1401 -----GCTTCAATGCTTACAGAGAGGCGCAAAATAGCCCACTTATGATGGCAACT 1454
Qy 288 ArgCysIysIleAlaGlnIleValAlaIleAsnGlyIleAsnPheLeuHisGluAsn----- 305
Db 1455 CGCAAGCGCATACATGAGGTTCTGCAAGAGGCTCTCCATCTTGATGATCATTTGTGAT 1514
Qy 306 ---HisIleIleHisArgAspIleIysSerAlaAsnIleLeuLeuAspGluAlaPheThr 324
Db 1515 CTTAAGATTATTTACCGGAGTGTCAAGGCTGTCAACATCTTACTGTGAATGAATATGAG 1574
Qy 325 AlaIysIleSerAspPheGlyLeuAlaArgAlaSerGluIysPheAlaGlnThrValMet 344
Db 1575 GCAGTGTGGGAGATTTTGGCTTGGCAAACTTATGAT---TATTAAGACACACATGTT 1631
Qy 345 ThrSerArgIleValGlyIleThrAlaIysMetAlaProGluAlaLeu---ArgIysGlu 363
Db 1632 ACAGACGCTGTCTGTGAACCATTTGACCATAGCACTGATGATGATCTTCTTACTGGAAG 1691
Qy 364 IleThrProIysSerAspIleTyrSerPheGlyValValLeuLeuGlnIleIleThrGly 383
Db 1692 TCTTGGAAAGACAGACGATTTGTGATATGATATGATGTGTGTGAACATTAATTACGGGA 1751
Qy 384 LeuProAlaValAspGluHisArg-----GluProGlnLeuLeuAsp--- 398
Db 1752 CAACGGGCAATTTTACCTTGCACCTTTAAGCAATGATGATGATGCTTGTGACTGG 1811
Qy 399 IleIysGlnIleGluAspGlnIleGluIysThrIleGluAspTyrIleAspIysMet 418
Db 1812 GTTAAAGCTTACTA-----AAAGAGAAAGCTTGAATGTGATGATGATCTTCTT 1865
Qy 419 ---AsnAspAlaAspSerThrSerValGluAlaMetCysIysSerValAlaSerGlnCysLeu 437
Db 1866 AAGAACATATATGTGAAGCAGAGGTGAACCACTTATCAAGTGTGATTAATTGTACA 1925

```

QY 438 HisGluLysAsnLysArgProAspIleLysValGlnGlnLeuGln 455
 DB 1926 CAAGGTACCAAGTATGATGACCAAGATGCTGAAGTGTAGATGTGAA 1979

RESULT 12
 US-08-447-185-2
 ; Sequence 2, Application US/08447185
 ; Patent No. 5648599
 ; GENERAL INFORMATION:
 ; APPLICANT: Tankeley, Steven D.
 ; APPLICANT: Martin, Gregory B.
 ; TITLE OF INVENTION: GENE CONFERRING DISEASE RESISTANCE
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Michael L. Goldman
 ; STREET: Clinton Square, P.O. Box 1051
 ; CITY: Rochester
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 14603
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/447,185
 ; FILING DATE:
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/111,078
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Goldman Mr., Michael L.
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (716)263-1000
 ; TELEFAX: (716)-263-1600
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 966 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-447-185-2

Alignment Scores:
 Pred. No.: 7.69e-41 Length: 966
 Score: 430.50 Matches: 114
 Percent Similarity: 49.55% Conservative: 52
 Best Local Similarity: 34.03% Mismatches: 100
 Query Match: 18.04% Indels: 69
 DB: 1 Gaps: 12

US-10-001-254-16 (1-460) x US-08-447-185-2 (1-966)

QY 141 LeuGlnLysSerIYrMetProPheArgSerSerProGluAsnLysSerLeuGlnVal 160
 DB 46 TTAAGCTTCGATGATCTCGTTCCTTTGAAAGT----- 78

QY 161 SerAspThrArgPheHisSerPheSerPheIYrGluLeuLysAsnValThrAsnAspHe 180
 DB 79 -----TATCGAGTTCCCTTTAGTATGTTGGAGGAACAACATAATATTTT 123

QY 181 AspGluArgProIleSerValGlyLysAsnLysMetGlyGlnGlyPheGlyValVal 200
 DB 124 GATCACAAGTTTATTA-----ATGCAACATGGTGTCTTTGGGAAGTT 165

QY 201 TyrLysGlyIYrVal---AsnAsnThrThrValAlaValLysLysLeuAlaAlaMetVal 219
 DB 166 TACAAGGGTGTGTTGGCTGATGAGCAAAAGTGGCCCTGAAAGG----- 210

QY 220 AspIleThrThrGlnGluLeuLysGln-----GlnPheAspGlnGlnLysVal 236
 DB 211 -----GCTACACCTGAGTCTCTCACAGATTTTGAAGAGTTTCAAAAGAAATTGAGACT 264

QY 237 MetAlaLysCysGlnHisGluAsnLeuValGluLeuGlnGlyPheSerSerAspGlyAsp 256
 DB 265 CTCTCATTTTGCAGACATCCGATCTGGTTTCATTGATAGGATTTCTGTGATGAAGAANAAT 324

QY 257 AspLeuCysLeuValIYrValIYrMetProAsnGlySerLeuLeuAspArgLeuSerCys 276
 DB 325 GAGATGATTTCAATTTAATTAATACATGAGATGGAAGCACTCAAGACATTTGTATGCA 384

QY 277 LeuAsp---GlyThrProPheSerThrPheIleMetArgCysLysIleAlaGlnGlyAla 295
 DB 385 TCAATCTACCCACCAAGAGCATGAGCTGGAGCAGACAGCTGGAGATTTGATGAGGCA 444

QY 296 AlaAsnGlyIleAsnPheLeuHisGluAsnHisIleHisArgAspIleLysSerAla 315
 DB 445 GCCAGAGGTCTACACTACCTTCATACAGACATTTATACATGATGATGATGATGATGAT 504

QY 316 AsnIleLeuLeuAspGluAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAla 335
 DB 505 AACATATTGCTTGATGAGAAATTTGTGCCAAAATTAATCTGATTTTGGAAATTCACAGAA 564

QY 336 SerGluLysPheAlaGlnThrValMetSerArgIleValGlyThrThrAlaIYrMet 355
 DB 565 GCGACTGAGCTTGATCAACCATCTTACAGCAGTACAGTGAAGAAAGCACTTCGGCTACATT 624

QY 356 AlaProGlu---AlaLeuArgGlyGluIleThrProLysSerAspIleYrSerPheGly 374
 DB 625 GACCTGATATTTTATTAAGAAGGAGCACTACCTGAAAATCTGATGATTTATCTTGGGT 684

QY 375 ValValLeuLeuGlu-----IleIleThrGlyLeuPro----- 385
 DB 685 GTTGTTTATTCGAAGTCTTTGTGCTAGTCTGCCATGATTCATCTCTTCCAAAGGAG 744

QY 386 -----AlaValAspGluHisArgGluProGlnIleLeuLeuAsp 398
 DB 745 ATGTTAATTTAGCTGATGAGGCGATGAGCGATGAATTAATTAAGACAGTTG----- 795

QY 399 IleLysGluLysLeuAspGluGlnLysThrIleGluAspTyrIleAspLysLysMet 418
 DB 796 -----GACCAATCTGATGATCC----- 813

QY 419 AsnAspAlaAspSerThrSerValGluAlaMetYr-----SerValAlaSerGln 435
 DB 814 AATCTTCAGATTAATAATTAAGACAGAGTCCCTCAGAAAGTTTGGAGATACAGCGGTAAA 873

QY 436 CysLeuHisGluLysLysAsnLysArgProAspIleLysVal 450
 DB 874 TGCTTACCTTGTCTGATGAGATGAGGCATCAATGAGGTGATGTTG 918

RESULT 13
 US-08-447-185-3
 ; Sequence 3, Application US/08447185
 ; Patent No. 5648599
 ; GENERAL INFORMATION:
 ; APPLICANT: Tankeley, Steven D.
 ; APPLICANT: Martin, Gregory B.
 ; TITLE OF INVENTION: GENE CONFERRING DISEASE RESISTANCE
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Michael L. Goldman
 ; STREET: Clinton Square, P.O. Box 1051
 ; CITY: Rochester
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 14603
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/447,185
 FILING DATE:
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/111,078
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Goldman Mr. Michael L.
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (716)263-1000
 TELEFAX: (716)-263-1000
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2443 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-08-447-185-3

Alignment Scores:
 Pred. No.: 3,46e-40 Length: 2443
 Score: 430.50 Matches: 114
 Percent Similarity: 49.55% Conservative: 52
 Best Local Similarity: 34.03% Mismatches: 100
 Query Match: 18.04% Indels: 69
 DB: Gaps: 12

US-10-001-254-16 (1-460) x US-08-447-185-3 (1-2443)

Qy 141 LeuGlulnserThrMetProAspSerSerProGluAsnLysSerLeuGluVal 160
 Db 109 TTAAGCTCGAGTAACTCGCTTCTTTGAAGT----- 141
 Qy 161 SerAspThrArgPheHisSerPheSerPheThyGluLeuLysAsnValThrAsnAspPhe 180
 Db 142 -----TATCGAGTTCCTTATGATATTGGAGGAGCAAGCACTAATAATTTT 186
 Qy 181 AspGluArgProIleSerValGlyGlyAsnLysMetGlyGlyGlyPheGlyValVal 200
 Db 187 GATCACAAGTTTAA-----ATTGACATCGGTCTTGGGAGAGCTT 228
 Qy 201 TyrIleGlyTyrVal-----AsnAsnThrThyValAlaValLysLysLeuAlaAlaMetVal 219
 Db 229 TACAAGGAGTGTTCCTGATGAGCAAGAGTGCCCTGAAAGG----- 273
 Qy 220 AspIleThrThrGluLysLeuLysGln-----GlnPheAspGlnLysVal 236
 Db 274 -----CGTACAGCTGAGTCTCTCAAGAGTATTGAAGAGTTCCGAACAGAAATTGAGACT 327
 Qy 237 MetAlaLysCysGlnHisGluAsnLysValGluLeuLysGlyPheSerSerAspGlyAsp 256
 Db 328 CTCCTCATTTGGACACATCCGCACTCTGTTTCATTGATAGCATTCGTGATGAAGAAT 387
 Qy 257 AspLeuLysLeuValTyrValTyrMetProAsnGlySerLeuLeuAspArgLysSerCys 276
 Db 388 GAGATGATTTCAATTATTAATACATGAGATGAGAACCTCAAGAGACATTGTATGGA 447
 Qy 277 LeuAsp---GlyThrProProLysSerThrHisMetArgCysLysIleAlaGlyVal 295
 Db 448 TCAGATTCACCCACATGATGATGAGCTGAGAGAGAGCTGGAGATHTGATAGGGGCA 507
 Qy 296 AlaAsnGlyIleAsnPheLysGluAsnHisIleHisIleHisIleHisIleHisIleHisIle 315
 Db 508 GCCAGAGGTCTACACTCTTACATCTAGACCAATTATACATCGATGCAAGTCTATA 567
 Qy 316 AsnIleLeuLeuAspGluAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAla 335
 Db 568 AACATATTGCTTGATGAGAAATTTGTGCCAAATAATCTGATTTTGGATATCCAGAAA 627

Qy 336 SerGluLysPheAlaGlnThrValMetThrSerArgIleValGlyThrThrAlaTyrMet 355
 Db 628 GGGAGCTGAGCTTATGATCAACCCCATCTTAGACAGTATGAAAGAACTCTCGGCTACATT 687
 Qy 356 AlaProGlu---AlaLeuArgGlyGlyIleThrProLysSerSerAspIleTyrSerPheGly 374
 Db 688 GACCCTGAAATATTTTAAAGGAGCAGCTCAGTGAATAATCTGATGTTATCTTCGCT 747
 Qy 375 ValValLeuLeuGlu-----IleIleThrGlyLeuPro----- 385
 Db 748 GTTCTTTATTCAGAGTCTTCTTGTCTAGGCTCGCATAGTTCAATCTCTCCAAAGGAG 807
 Qy 386 -----AlaValAspGluHisArgGluProGlnLeuLeuAsp 398
 Db 808 ATGTTAATTAGCTGAATGAGCAGTGGAGTCCGATTAATATGACAGATTG----- 858
 Qy 399 IleGluGluLysIleGluAspGluLysThrIleGluAspTyrIleAspLysLysMet 418
 Db 859 -----GAAACAATCGTATGCC----- 876
 Qy 419 AsnAspAlaAspSerThrSerValGluAlaMetTyr-----SerValAlaSerGln 435
 Db 877 AATCTTGACATTAATAATAGACAGAGTCCCTCAGAGAGTTGGAGATACAGCGCTAA 936
 Qy 436 CysLeuHisGluLysLysAsnLysArgProAspIleLysLysVal 450
 Db 937 TGCTTAGCTTTGCTAGTGAAGATGAGCCATCAATGAGTGTG 981

RESULT 14

US-07-717-331F-4
 Sequence 4, Application US/07717331F
 Patent No. 5484905

GENERAL INFORMATION:

APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
 TITLE OF INVENTION: A Receptor Protein Kinase Gene
 TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Yahwak & Associates
 STREET: 25 Skytop Drive
 CITY: Trumbull
 STATE: Connecticut
 COUNTRY: USA
 ZIP: 06611
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy Disk
 COMPUTER: Macintosh
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: Microsoft Word
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/717,331F
 FILING DATE: June 19th 1991
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: George M. Yahwak
 REGISTRATION NUMBER: 26,824
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (203)268-1951
 TELEFAX: (203)268-1951
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2749 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA
 US-07-717-331F-4

Alignment Scores:
 Pred. No.: 9,43e-40 Length: 2749
 Score: 427.50 Matches: 152
 Percent Similarity: 44.25% Conservative: 75

; LENGTH: 3459 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 34..1908
 ; US-08-980-060-3

Alignment Scores:

Pred. No.: 3.53e-39 Length: 3459
 Score: 424.00 Matches: 121
 Percent Similarity: 46.19% Conservative: 97
 Best Local Similarity: 25.64% Mismatches: 172
 Query Match: 17.77% Indels: 82
 DB: Gaps: 14

US-10-001-254-16 (1-460) x US-08-980-060-3 (1-3459)

QY 53 ILeaRgArPheGluAlaLeuLeuGlnThr---GlyYsSerProThrSerGluLeuLeu 71
 Db 157 CTGGGAAGATCAAGTCATGAGCGGGTGCAGGGTGTGAGCATCACCGGAGACTGCTG 216
 QY 72 PheAspTrpGlyThrThrAsnCysThrValGlyAspLeuValAspLeuLeuIleGlnAsn 91
 Db 217 TGGTGTGGGGCATGCGGACGCGCACCGTCCAGCAACTTGTGACCTCTGTGCGGCTG 276
 QY 92 GluPhePheAlaProAlaSerLeuLeu----- 101
 Db 277 GAGCTTACCGGGCTGCTCCAGATCATCTGAACTGAAACCGGCTCTGAAATCAGGTGT 336
 QY 102 -----ProAspAlaVal---ProLYsThrAlaAsnThrLeuProSerLYs 115
 Db 337 CCCATTCCAGGCTTCCTGACTCTGTGAGCCAGCAAAAGCTTTGGCAGCTTCTGTAGA 396
 QY 116 GluAlaIleThrValGlnGlnLYsGlnMetProPheCysAspLYsAspArgThrLeuMet 135
 Db 397 AAGGTGAGATGAACAGAAAGAGGGGACGCTGTG---AGAGTGGCCACTTTTCCA 450
 QY 136 ThrProValGlnAsnLeuGlnGlnSerLYr-----MetProProAspSer 150
 Db 451 GGGCCAGGCTCTCTCCAGCCAGCCACGACGCGCTTTCTCCAGCTCTCGAAGAA 510
 QY 151 SerSerProGluAsnLYsSerLeuGluVal-----SerAspThrArgPheHisSer 167
 Db 511 GATGCCCTCTCTCTTCTTGAGAGGACCTCCCACTTGCTGTGATTCAAAGACTTCAAC 570
 QY 168 -----PheSer 169
 Db 571 ACCTCCATTCTTAAGAGAAAACTTTGAGCTTGCTGGAGACAGCCTTTTCTGGAGT 630
 QY 170 PheTYrGluLeuLYsAsnValThrAsnAsnPheAspGluArgProIleSerValGlyGly 189
 Db 631 GAGGCAACGCTGTCAGGCAACCGATGACTTCATCA-----AAC 672
 QY 190 AsnLYsMetGlyGlnGlyLYsPheGlyValValTYrLYsGlyTYrValAsnAsnThrThr 209
 Db 673 CGCAAAATCAGCCAGGGGACCTTTGCTGACGCTTACAGAGGACAGGACGCGGAGCCA 732
 QY 210 ValAlaValLYsLYsLeuAlaIleMetValAspIleThrGlnGluLeuLYsGlnGln 229
 Db 733 TTGCTTTCAAGAGCTCAGAGACAGCCTGTTCAAGTCAGGATCAATGAAAGATTC 792
 QY 230 PheAspGlnGluIleLYsValMetAlaLYsCysGlnHisGluAsnLeuValGluLeu 249
 Db 793 TTCAGGACAGATGTCAGATTGTCTTAGATGTCGCCACCCCAATGCTTACTGCTGCTG 852
 QY 250 GlyPheSerSerAspLYsAspAspLeuLYsLeuValTYrValTYrMetProAsnGlySer 269
 Db 853 GGGCTTGTGTGTGCAAGACAGTTTCAAGCTTCACTACCCCTACATGCGCAATGAGTTC 912
 QY 270 LeuLeuAspArgLeuSerCysLeuAspGlyThrProProLeuSerTYrHisMetArgCys 289

Db 913 CTACAGGACAGACTGCAAGGCTCAGGGTGGTGGCTCGAGCCCTCCCTCCGCGCCAGCGTTC 972
 QY 290 LysIleAlaGlnGlyAlaIleAsnGlyIleAsnPheLeuHisGluAsnHisIleHis 309
 Db 973 AGCATCTGCTCAGAGGCTGCTCTGTGCCGTGAGTACTGATGCTGTGAGATCATTCAC 1032
 QY 310 ArgAspIleLYsSerAlaAsnIleLeuLeuAspGluAlaPheThrAlaLYsIleSerAsp 329
 Db 1033 AGCAAGTCAAAGGCTTAAATGTCTGTGACCAAAATCTCACCCCAACTTGTCTCAC 1092
 QY 330 -----PheGlyLeuAlaArgAlaSerGluLYsPheAlaGlnThrValMetThrSerArg 347
 Db 1093 CCNATGGCTCATGTGTCTCTGTCAACAAAGGTCAAAATACCAATGATGAAGACTCAC 1152
 QY 348 ILeValGlyThrThr---AlaTYrMetAlaProGluAlaLeuArg---GlyGluIleThr 365
 Db 1153 CTGCTCCGACGTGACAGCGCGATCTGCCCAGAGATTTCACTCGGGTGGGCGAGTACCA 1212
 QY 366 ProLYsSerAspIleTYrSerPheGlyValValLeuLeuGlnIleIleThrGlyLeuPro 385
 Db 1213 AAGGAGTGGACATCTTCACTGAGATATGTGTGGCCGAGGTCTCTACAGGCGATCCCT 1272
 QY 386 AlaValAspGluHisArgGluProGlnLeuLeuAspIle----- 399
 Db 1273 GCATGTGATTAACAACGAAACCGCGGTTTAACTGAAGACTTACTCTCAGTGAATTCCA 1332
 QY 400 -----LYsGluGluIleGluAsp-----GluGluLYs 408
 Db 1333 AGCAACACCGCTCGCTCTGCTCCAGAAACAGCGGGGTGAGAACTGATGCGAAAGAG 1392
 QY 409 ThrIleGluAspTYrIleAspLYsLYsMetAsnAspAlaAspSerThrSerValGluAla 428
 Db 1393 ATGTCGACAGAAATCACTGAGAAAGGCGCAGGAGCTTCCGAGAGACTCGCCGAGGCC 1452
 QY 429 MetTYrSerValAlaSerGlnCysLeuHisGluLYsIleAsnLYsArgProAspIleLYs 448
 Db 1453 CTGGCCACGCGCTGCTGCTGTGCTGTCGCGGAGCGCTGAACCCAGC----- 1497
 QY 449 LysValGlnGlnLeuGlnGlnMetThrAlaSer 460
 Db 1498 -----CTGCAGAGAGTGTGTGCTCT 1518

Search completed: January 18, 2004, 06:10:41
 Job time : 145.032 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 18, 2004, 04:27:15 ; Search time 688.136 Seconds
(without alignments)
2356.207 Million cell updates/sec

Title: US-10-001-254-16
Perfect score: 2386
Sequence: 1 MNKEITPSTYVRCINVGILR.....KKKBPDIKKVOQLIQEMTAS 460

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2324096 seqs, 1762381658 residues
Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPRO.spool/US10001254/runat_16012004_152426_19815/app.query.fasta_1.1109
-DB=Published Applications_NA -qfmt=fastcap -SUFFIX=trmb -MINMATCH=0.1
-LOOFCU=0 -LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=D1osum62
-TRANS=numa40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10001254 @CGN 1.1 534 @runat_16012004_152426_19815
-NCPU=6 -ICU=3 -NO_MMAR -LARGEOBERT -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOF=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/FCI_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
------------	-------	--------------------	----	-------------

1	2386	100.0	1383	15	US-10-001-254-15	Sequence 15, Appl
2	2382	99.8	1383	11	US-09-759-595-2	Sequence 2, Appl
3	2362	99.0	2817	11	US-09-966-451-3	Sequence 3, Appl
4	2362	99.0	2817	15	US-10-001-254-27	Sequence 27, Appl
5	2017	84.5	1542	11	US-09-759-595-4	Sequence 4, Appl
6	972	40.7	1493	10	US-09-764-668-249	Sequence 249, App
7	779	32.6	483	9	US-09-745-763-109	Sequence 109, App
8	691	29.0	501	9	US-09-833-790-149	Sequence 149, App
9	590	24.7	31000	11	US-09-966-451-10	Sequence 10, Appl
10	517	21.7	294	15	US-10-001-254-5	Sequence 5, Appl
11	515.5	21.6	3590	10	US-09-880-107-2222	Sequence 2222, Ap
12	515.5	21.6	3590	12	US-10-167-034-4	Sequence 4, Appl
13	515.5	21.6	3590	13	US-10-101-510-88	Sequence 88, Appl
14	515.5	21.6	3590	13	US-09-873-667C-1003	Sequence 1003, Ap
15	508.5	21.3	3261	10	US-09-938-842A-1673	Sequence 1673, Ap
16	506	21.2	1899	10	US-09-938-842A-799	Sequence 799, App
17	499	20.9	1281	10	US-09-938-842A-624	Sequence 624, App
18	498.5	20.9	2607	10	US-09-938-842A-1269	Sequence 1269, Ap
19	494	20.7	1389	10	US-09-938-842A-340	Sequence 340, App
20	485	20.3	1239	10	US-09-938-842A-1190	Sequence 1190, Ap
21	481	20.2	2196	14	US-10-086-464-16	Sequence 16, Appl
22	481	20.2	2261	14	US-10-086-464-15	Sequence 15, Appl
23	480.5	20.1	2640	10	US-09-938-842A-1270	Sequence 1270, Ap
24	479	20.1	1275	10	US-09-938-842A-1977	Sequence 1977, Ap
25	478.5	20.1	1944	14	US-10-086-464-1	Sequence 1, Appl
26	478.5	20.1	2189	14	US-10-086-464-3	Sequence 3, Appl
27	477.5	19.8	2733	15	US-10-101-664A-869	Sequence 869, App
28	477.5	19.8	3257	15	US-10-101-664A-455	Sequence 455, App
29	472.5	19.8	3257	15	US-10-101-664A-456	Sequence 456, Appl
30	471	19.7	1257	13	US-10-342-224-67	Sequence 67, Appl
31	470	19.7	3453	13	US-10-259-165-311	Sequence 311, App
32	468.5	19.6	2690	15	US-10-101-664A-870	Sequence 870, App
33	466.5	19.6	2466	10	US-09-938-842A-2644	Sequence 2644, Ap
34	460	19.3	1098	10	US-09-938-842A-2233	Sequence 2233, Ap
35	459.5	19.3	3237	15	US-10-101-664A-858	Sequence 858, App
36	457.5	19.2	1419	10	US-09-938-842A-926	Sequence 926, App
37	456	19.1	2124	10	US-09-938-842A-1079	Sequence 1079, Ap
38	456	19.1	2242	13	US-10-304-946-7	Sequence 7, Appl
39	456	19.1	2144	13	US-10-304-946-6	Sequence 6, Appl
40	456	19.1	2411	13	US-10-304-946-8	Sequence 8, Appl
41	456	19.1	2736	10	US-09-938-842A-56	Sequence 56, Appl
42	455	19.1	70768	14	US-10-135-322-13	Sequence 13, Appl
43	455	19.1	2685	11	US-09-754-853A-7	Sequence 7, Appl
44	455	19.1	2685	11	US-09-754-853A-50	Sequence 50, Appl
45	455	19.1	2685	11	US-09-754-853A-51	Sequence 51, Appl

ALIGNMENTS

RESULT 1
US-10-001-254-15
; Sequence 15, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sung Hyung
; APPLICANT: Roch, Wilfried
; APPLICANT: Steiner-Liwen, Frank
; TITLE OF INVENTION: No. US20030049702A1e1 Death Domain Proteins
; FILE REFERENCE: P-LJ 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1383

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1383)
US-10-001-254-15

```

Alignment Scores:

```

Pred. No.: 3,756-259 Length: 1383
Score: 2386.00 Matches: 460
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

```

US-10-001-254-16 (1-460) x US-10-001-254-15 (1-1383)

```

QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
DB 1 ATGAACAACCCATTAACCATCAATATGCGCTGCTCAATGTTGACTAATTAGG 60
QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPheValAlaIleLeu 40
DB 61 AAGCTGCAGATTTTATTTGATCCCAAGAGAGATGGAAGATTAGCTATTAGAAA 120
QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgAspPheGluAlaLeu 60
DB 121 AATCACTCTGGTATGATGATCAATCAATCACTTCAATAGGAATTTGAAGCATTTACTT 180
QY 61 GlnThrGlyLysSerProThrSerGlnLeuLeuPheAspTyrGlyThrAsnCysThr 80
DB 181 CAATCTGGAAGAAATGCCCACTTCTGAATTACTGTTTGACTGGGACACCAATTTGCACA 240
QY 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPheAlaProAlaSerLeu 100
DB 241 GTTGATGATCTTGATGATCTTTATCAAAATGATTTTGTCTCCGACAGCTTTTG 300
QY 101 LeuProAspAlaValProLysThrAlaAsnThrLeuProSerLysGluAlaIleThrVal 120
DB 301 CTCACCATGCTGTTCCCAAAATGCTAATCACTACTCTTAAAGAACTAATACAGTT 360
QY 121 GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140
DB 361 CAGCAAAAACAGATGCTTTCTGTGACAAAGACAGACATGATGACCTGTGCAGAT 420
QY 141 LeuGlnGlnSerTyrMetProProAspSerSerSerProGluAsnLysSerLeuGluVal 160
DB 421 CTTGAACAAAGCTATATGCCACTGACTCTCAAGTCCAGAAAATTAAGTTTGAAGTT 480
QY 161 SerAspThrArgPheHisSerPheSerPheTyrGlnLeuLysAsnValThrAsnAspPhe 180
DB 481 AGTATATACCTTTTTCACAGTTTTCATTTATATGATTTGAAGATGACAAATTAACCTT 540
QY 181 AspGlnArgProIleSerValGlyLysAsnLysMetGlyGlnGlyPheGlyValVal 200
DB 541 GATGAACGACCACTTTCTGTTGTTGTTGTTAATAATGGAGAGGAGATTGGAGTTGTA 600
QY 201 TyrTyrGlyTyrValAsnAsnThrThrValAlaValLysLysLeuAlaIleMetValAsp 220
DB 601 TATTAAGGCTACGTAAATTAACAACTGTGACAGTGAAGAACTTGCAGCAATGTTGAC 660
QY 221 IleThrThrGlnGluLeuLysGlnPheAspGlnGluIleLysValMetAlaLysCys 240
DB 661 ATTATCTACTGAAGAACTGAACAGCACTTTATCAAGAAATTAAGTATATGCAAAAGTGT 720
QY 241 GlnHisGluAsnLeuValGluLeuLeuGlyPheSerSerAspGlyAspAspLeuCysLeu 260
DB 721 CAACATGAAACCTTAGTAGAATCTTCTGTTCTCAAGATATGAGATGACCTCGCTTA 780
QY 261 ValTyrValTyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThr 280
DB 781 GTATATGTATTACATGCTTAATGTTCAATGCTAGACACACTCTCTGCTGATGATGACT 840

```

```

QY 281 ProProLeuSerTrpHisMetArgCysLysIleAlaGlnGlyValAlaAsnGlyIleAsn 300
DB 841 CCACCACTTTCTTGGCACATGATGACAGATTGCTCAGGGTGCAGCTAATGCACTCAAT 900
QY 301 PheLeuHisGluAsnHisHisIleHisArgAspIleLysSerAlaAsnIleLeuLeuAsp 320
DB 901 TTCTCATGATAAATCATCATATTTATGAGATATTAAGTCAATAATCATTAACGAT 960
QY 321 GluAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGluLysPheAla 340
DB 961 GAAGCTTTTCTCTTAATATCTGACTTTGCTGACGGCTTGCAGGGCTCTGAGAACTTGGC 1020
QY 341 GlnThrValMetThrSerArgIleValGlyThrThrAlaTyrMetAlaProGluAlaLeu 360
DB 1021 CAGACGTATGATGATGACAAATTTGGGAAACAACAGCTTATATGACCAAGACTTTTG 1080
QY 361 ArgGlyGluIleThrProLysSerAspIleTyrSerPheGlyValValLeuLeuGluIle 380
DB 1081 CGTGGAGAAATTAACCCAAATCTGATATTACAGCTTGTGTGTTTACTAGAAATA 1140
QY 381 IleThrGlyLeuProAlaValAspGluHisArgGluProGlnLeuLeuAspIleLys 400
DB 1141 ATTACTGACTTCAGCTGTGATGATGAACCGTAACCTCACTTATTTCTAGATATTAAA 1200
QY 401 GlnGluIleGluAspGlnGluLysThrIleGluAspTyrIleAspLysMetAsnAsp 420
DB 1201 GAGGAATTAAGATGATGAAGAAAGACATTTGAAGATTATATGATTAAGAAATGATAT 1260
QY 421 AlaAspSerThrSerValGluAlaMetTyrSerValAlaSerGlnCysLeuHisGluLys 440
DB 1261 GCTGATTCCTCCTTCACTGATGAAGCTATGATCTCTGTTGCTAGTCAATGCTGCAAGAA 1320
QY 441 LysAsnLysArgProAspIleLysLysValGlnGlnLeuLeuGlnGlnMetThrAlaSer 460
DB 1321 AAAAATTAAGACACACATTAAGAAAGTTCAACAGCTGCTGCAAGAGATGACACTTCT 1380

```

RESULT 2

```

US-09-759-595-2
; Sequence 2, Application US/09759595
; Publication No. US20030059916A1
GENERAL INFORMATION:
; APPLICANT: Wesche, Holger
; APPLICANT: Li, Shyun
; APPLICANT: Tularik Inc.
TITLE OR INVENTION: IRAK-4: Compositions and Methods of Use
FILE REFERENCE: 018781-003910US*
CURRENT APPLICATION NUMBER: US/09/759,595
PRIOR FILING DATE: 2001-01-13
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1383
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human IL-1 receptor-associated kinase 4 (IRAK-4)
OTHER INFORMATION: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1383)
OTHER INFORMATION: human IRAK-4
US-09-759-595-2

```

Alignment Scores:

```

Pred. No.: 1,066-258 Length: 1383
Score: 2382.00 Matches: 459
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.83% Indels: 0
Gaps: 0

```

US-10-001-254-16 (1-460) x US-09-759-595-2 (1-1383)

QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
 Db 1 ATGAACAAACCCATACACCATCAATATGTGCGCTCCCTCAATGTGGACTATATGG 60
 QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTPlsLysLeuAlaValAlaIleLys 40
 Db 61 AAGCTGCAGATTTTATGATCCTCAAGAGAGATGAAGAAGTTAGCTGTAGCTATTA 120
 QY 41 LysProSerGlyAspAspArgTyArgGlnPheHisIleArgArgPheGluAlaLeu 60
 Db 121 AAACCATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 61 GlnThrGlyLysSerProThrSerGlnLeuLeuPheAspTlPglYThrThraAsnCysThr 80
 Db 181 CMACTGAAACCAAGTCCCACTTGTGATGATGATGATGATGATGATGATGATGAT 240
 QY 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPheAlaProAlaSerLeuLeu 100
 Db 241 GCTGATGATCTTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGT 300
 QY 101 LeuProAspAlaValProLysThrAlaAsnThrLeuProSerLysGluAlaIleThrVal 120
 Db 301 CTCCAGATGCTGTGCTCCCAAACTGCTATATCACTTCTTAAAGAGCTATTAACGTT 360
 QY 121 GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140
 Db 361 CAGCAAAACAGATGCTTCTGTGACAAAGACGAGATGATGACCTGTGCAGAT 420
 QY 141 LeuGlnGlnSerTyMetProProAspSerSerSerProGlnAsnLysSerLeuGlnVal 160
 Db 421 CTGAAACCAAGCTATATGCTCACTGACTCTCAAGTCCAGAAATTAAGTTGAAGTT 480
 QY 161 SerAspThrArgPheHisSerPheSerPheTyGlnLeuLysAsnValThrAsnAspPhe 180
 Db 481 AGTATACACGTTTTCACAGTTTTCATTTATGATTAAGAAATGTCACAAATACATTT 540
 QY 181 AspGlnArgProIleSerValGlyLysAsnLysMetGlyGlnGlyPheGlyValVal 200
 Db 541 GATTAACGACCATTTCTGTGTGTGATTAATAAATGGAGAGGAGATTTGGAGTTGA 600
 QY 201 TyrlsGlyTyValAsnAsnThrThrValAlaValLysLysLeuAlaAlaMetValAsp 220
 Db 601 TATAAGGCTACGTAATTAACAACTGTGCACTGGAAGAGCTTGCACGATGCTTGC 660
 QY 221 IleThrThrGlnGlnLeuLysGlnGlnPheAspGlnGlnLysValMetAlaLysCys 240
 Db 661 ATTACTACTGAAGAACTGAACACAGATTTGATCAAGAAATTAAGTATGCAAGATGT 720
 QY 241 GlnHisGlnLeuLeuValGlnLeuLeuGlyPheSerSerAspGlyAspAspLeuCysLeu 260
 Db 721 CAACATGAACAACTTAAGTAAGAACTCTGTTCTCAAGTATGAGATGATGATGATGAT 780
 QY 261 ValTyValTyMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThr 280
 Db 781 GTATATGTTTACATGCTTAATGCTTCAATGCTTACACACATCTCTGCTTGGATGCT 840
 QY 281 ProProLeuSerTlPglIleMetArgCysLysIleAlaGlnGlyAlaAlaAsnGlyLysAsn 300
 Db 841 CCACACCTTCTTCTGCAATGATGCAAGATGCTGAGGCTGCAAGCTTAATGCAATCAT 900
 QY 301 PheLeuHisGlnLeuAsnHisIleHisArgAspIleLysSerAlaAsnIleLeuLeuAsp 320
 Db 901 TTTTACATGAACAACTCATATTCATAGATATTAAGTGAAGTGAATTCCTTAATCTGAT 960
 QY 321 GluAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGlyLysPheAla 340
 Db 961 GAACCTTTTACGCTTAATAATATGACCTTGGCTTGCACGCGCTTCTGAGAACTTTGCC 1020
 QY 341 GlnThrValMetThrSerArgIleValGlyThrThrAlaTyMetAlaProGlnAlaLeu 360
 Db 1021 CAGACAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080

QY 361 ArgGlyGlnIleThrProLysSerAspIleTySerPheGlyValLeuLeuGlnIle 380
 Db 1081 CGTGAACAAATTAACACCAATCTGATATTTACAGCTTTGTGTGTGTTTACTAGAAATA 1140
 QY 381 IleThrGlyLeuProAlaValAspGlnHisArgGlnProGlnLeuLeuAspIleLys 400
 Db 1141 ATACTGACCTTCAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 QY 401 GlnGlnIleGlnAspGlnGlnGlyThrIleGlnAspTyrlleAspLysLysMetAsnAsp 420
 Db 1201 GAAGAAATTCAGATGAGAAAGAAAGCAATGAAATTAATTAATAAAGATCAATGAT 1260
 QY 421 AlaAspSerThrSerValGlnAlaMetTySerValAlaSerGlnCysLeuHisGlnLys 440
 Db 1261 GCTGATTCACCTTCAGTGAAGCTATGATCTGTGTGATGATGATGATGATGATGAT 1320
 QY 441 LysAsnLysArgProAspIleLysLysValGlnGlnLeuLeuGlnGlnMetThrAlaSer 460
 Db 1321 AAAAATAGAGACCAACATTAAGAAGTTCAACAGCTGTGCAAGAGATGACAGCTTCT 1380

RESULT 3
 US-09-966-451-3
 ; Sequence 3, Application US/09966451
 ; Publication No. US2003008756A1
 ; GENERAL INFORMATION:
 ; APPLICANT: C. Frank Bennett
 ; APPLICANT: Susan W. Freier
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF IL-1 RECEPTOR-ASSOCIATED KINASE-4 EXPRES-
 ; FILE REFERENCE: RFS-0324
 ; CURRENT APPLICATION NUMBER: US/09/966,451
 ; CURRENT FILING DATE: 2001-09-28
 ; NUMBER OF SEQ ID NOS: 88
 ; SEQ ID NO 3
 ; LENGTH: 2817
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (50)...(1432)
 US-09-966-451-3

Alignment Scores:
 Pred. No.: 5 62e-256 Length: 2817
 Score: 2362.00 Matches: 456
 Percent Similarity: 99.13% Conservative: 0
 Best Local Similarity: 99.13% Mismatches: 4
 Query Match: 98.99% Indels: 0
 DB: 11 Gaps: 0

US-10-001-254-16 (1-460) x US-09-966-451-3 (1-2817)
 QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
 Db 50 ATGAACAAACCCATACACCATCAATATGTGCGCTCCCTCAATGTGGACTATATGG 109
 QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTPlsLysLeuAlaValAlaIleLys 40
 Db 110 AACCTGTGATTTTATGATCTTCAAGAGAGATGAGAGAGTATGATGATGATGAT 169
 QY 41 LysProSerGlyAspAspArgTyArgGlnPheHisIleArgArgPheGluAlaLeu 60
 Db 170 AAACCATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 229
 QY 61 GlnThrGlyLysSerProThrSerGlnLeuLeuPheAspTlPglYThrThraAsnCysThr 80
 Db 230 CAAACTGGAAGAAAGTCCCACTTCTGATTAATGCTTGTGACGTGGGACCAAAATTCAC 289
 QY 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPheAlaProAlaSerLeuLeu 100
 Db 290 GTTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGTG 349
 QY 101 LeuProAspAlaValProLysThrAlaAsnThrLeuProSerLysGlnAlaIleThrVal 120

```

Db      350 CTCCTCCGATGCTCTCTCCCAAACTGCTAAATACACTCTTAAGAGAGCTATTAACAGTT 409
Qy      121 GInGInLySGLInMeCProPheCyAsPlyAsPArGrThrLeuMeCThrProValGInAsn 140
Db      410 CAGCAAAAACAGATGCTTCTTGTCGACAAAGACAGGACATTTGATGACACTTGCGAGMAT 469
Qy      141 LeuGInGInSerLyMetProPAsPSeSerSerProGInuSnLySLeuGInuVal 160
Db      470 CTTGAAACNAGCTATATGCACTGACTCTCTCAAGTCCAAATAATTAAGTTTAAAGATT 529
Qy      161 SerAsPThrArgPheHisSerPheSerPheTyGInuLeuLyAsnValThrAsnAsnPe 180
Db      530 AGTGAATACAGTTTTCACAGTTTTCATTTTAATTAAGATGACCAATAACTTT 589
Qy      181 AsPGLuArProLieserValGlyGlyAsnLyMeCGLyGInuGlyGlyPheGlyValVal 200
Db      590 GATGAACGACCCATTTCTGTTGGTGATTAATAATGGAGAGAGGATTTGAGATTGTA 649
Qy      201 TyfLySGLyTyfValAsnAsnThrThrValAlaValLySLeuAlaAlaMetValAsP 220
Db      650 TATTAAGGCTAAGTAATTAACAATGCTGCGAGTGAAGAGCTTGCAAGCAATGCTTGAC 709
Qy      221 IleThrThrGInuGInuLeuLySGLInPheAsPGLInuLySValMetAlaLyS 240
Db      710 ATTACTACTGAGAACTGAAACAGCACTTTGATCAAGAAATTAAGTAATGGCAAGTGT 769
Qy      241 GInHiSGLuAsnLeuValGInuLeuGlyPheSerSerAsPGLyAsPAsPLeuCySLeu 260
Db      770 CAACATGAAAACTAGTAGAACTACTGTTTCTCAAGTATGAGATGAGATGACCTCTGCTTA 829
Qy      261 ValTyfValTyfMetProAsnGlySerLeuLeuAsPArgPheSerCySLeuAsPGLyThr 280
Db      830 GTTATATGTTTACATGCTTAATGTTTCTTCTGACACAGACTCTCTTGCTTGAGTGTACT 889
Qy      281 ProProLeuSerTrpHisMeCArgCySLeuIleAlaGInuLyAlaAlaAsnGlyIleAsn 300
Db      890 CCAACACTTCTTGCGACATGAGATGCAAGATTGCTCAGGGTGCAGCTAATGAGATCAAT 949
Qy      301 PheLeuHiSGLuAsnHisHisIleHisArgAsPGLySLeuSerAlaAsnIleLeuLeuAsP 320
Db      950 TTTCTCATGAAAAATCATCATATTCATATGAGATTAATAAGTGAATAATCTTAACTGAT 1009
Qy      321 GluAlaPheThrAlaValIleSerAsPArgPheGlyLeuAlaAlaProLAserGInuLyS 340
Db      1010 GAAGCTTTTACTGCTAATAATATCTGACTTTGGCTTGACGGCTTCTGAGAAAGTTGCC 1069
Qy      341 GInThrValMetThrSerArgIleValGlyThrThrAlaTyfMetAlaProGInuAlaLeu 360
Db      1070 CAGCAGATCATGACTAGCAGAAATGTTGGAAACAACAGCTTATATATGCAACAGAAAGCTTTG 1129
Qy      361 ArgGlyGInuLeuThrProLySerAsPGLyTyfSerPheGlyValValIleLeuGInuLeu 380
Db      1130 CGTGGAAATAATACACCCCAATCTGATATTAACGCTTGTGTGCTTTTATCAAGAAATA 1189
Qy      381 IleThrGlyLeuProAlaValAsPGLuHisArgGInuProGInuLeuLeuAsPGLyLeu 400
Db      1190 ATATCTGAGCTTCCAGCTGTGAGTGAACACCGTGAACCTCACTATATGCTAGATATTAATA 1249
Qy      401 GInuGInuLeuGInuAsPGLuLySThrIleGInuAsPGLyIleAsPGLySLeuAsnAsP 420
Db      1250 GAAGAATTTGAAGTGAAGAAAGACAAATGGAATTAATTAAGTAAGTAAGTAAGTAAGT 1309
Qy      421 AlaAsPSeThrSerValGInuAlaMetTyfSerValAlaSerGInuCySLeuHisSGLuLyS 440
Db      1310 GCTGATTTCCACTTCACTTGAAGCTATGTACTGTGGTCTAGCCATATGCTCGGCAATGAAGA 1369
Qy      441 LyAsnLySArgProAsPGLyLeuLySValGInuGInuLeuGInuGInuMetThrAlaSer 460
Db      1370 AAAAATAAGAGCCAGACATTAAGAGGTTCAACAGCTGCTGCAAGAGATGACAGCTTCT 1429

```

RESULT 4
US-10-001-254-27

```

; Sequence 27, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sung Hyung
; APPLICANT: Roth, Wilfried
; APPLICANT: Stenmer-Liwen, Frank
; TITLE OF INVENTION: No. US20030049702A1e1 Death Domain Proteins
; FILE REFERENCE: P-Ld 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 2817
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50) ... (1429)
; US-10-001-254-27

Alignment Scores:
Pred. No.: 5,62e-256 Length: 2817
Score: 2362.00 Matches: 466
Percent Similarity: 99.13% Conservative: 0
Bee Local Similarity: 99.13% Mismatches: 4
Query Match: 98.99% Indels: 0
DB: 15 Gaps: 0

US-10-001-254-16 (1-460) x US-10-001-254-27 (1-2817)

Qy      1 MetAsnLySProLLeuThrProSerThrTyfValAlaGySLeuAsnValGlyLeuIleArg 20
Db      50 ATGAACAAACCCATTAACACATCAACATATGCGCGCTCAAGTTGAGCTTAATTAAG 109
Qy      21 LySLeuSerAsPArgPheHisSerProGInuGInuLySLeuAlaAlaIleLyS 40
Db      110 AACCTTCAGATTTTATATCTCTCAAGAAAGATGAGAAAGTTAGCTTACTATTAATA 169
Qy      41 LySProSerGlyAsPAsPArgTyfAsnGInuPheHisIleArgArgPheGInuAlaLeu 60
Db      170 AAACCATCTGGATGATGATATGATACATCACTTTCATATGAGAGATTGAACATTAATT 229
Qy      61 GInThrGlyLySLeuProThrSerGInuLeuPheAsPGLyThrThrAsnCySThr 80
Db      230 CAAACGAGAAAAAGTCCCACTTCTGAAATTAAGTTGACTGGGGCCCAACAATTTGCA 289
Qy      81 ValGlyAsPLeuValAsPLeuLeuIleGInuGInuPheAlaProLAserLeu 100
Db      290 GTTGGATCTGTGGATCTTTGATCCAAATGAATTTTGTCTCTCGAGTCTTTTG 349
Qy      101 LeuProAsPAlaValProLySThrAlaAsnThrLeuProSerLySGLuAlaIleThrVal 120
Db      350 CTCCTCCGATGCTCTCTCCCAAACTGCTAATATACACTTCTTAAGAAAGCTATTAACAGTT 409
Qy      121 GInGInLySGLInMeCProPheCyAsPlyAsPArGrThrLeuMeCThrProValGInAsn 140
Db      410 CAGCAAAAACAGATGCTTCTTGTCGACAAAGACAGGACATTTGATGACACTTGCGAGMAT 469
Qy      141 LeuGInGInSerLyMetProPAsPSeSerSerProGInuSnLySLeuGInuVal 160
Db      470 CTTGAAACNAGCTATATGCACTGACTCTCTCAAGTCCAAATAATTAAGTTTAAAGATT 529
Qy      161 SerAsPThrArgPheHisSerPheSerPheTyfGInuLeuLyAsnValThrAsnAsnPe 180

```

530 AGGTATACAGCGTTTTCACAGTTTTCATTTATGATTAAGAAATGCAAAATTAATCTT 589
Qy AepgluaqproilseervaiGlyYasnlysmetGlyuglyGlyPheGlyValVal 200
181
Db GATGAGACACCATTTCTGTTGGTGTATTAATGGAGAGAGGAGATTGGAGTTGTA 649
Qy TyrlysglyYrValAsnAsnThrThrValAlaValLysLysLeuAlaAlaMetValAsp 220
201
Db TATAAAGGCTACGTAAATATACACACTGTGGCAGTGAAGAACTTGCAGCAATGGTTAC 709
Qy 11ethrthrglugluLeuLysGlnGlnPheAspGlnGluLysValMetAlaLysCys 240
221
Db ATTACTACTGAAAGAACTGAAACAGCACTTTGATCAAGAAATTAAGATATGCAAAAGTGT 769
Qy 241 Glnh1sgluAsnLeuValGluLeuLeuGlyPheSerSerAspGlyAspAspLeuCysLeu 260
770 CACACGTAAACCTTAGTGAACCTACTGTTCTCAAGTGAAGAGATGACCTTGGCTTA 829
Qy 261 ValYrValYrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThr 280
830 GTATATGTTTACATGCCCTAATGTTTCATGCTAGACACACTCTGCTTGGATGTAATCT 889
Qy 281 ProProLeuSerThrPheMetArgCysLysLysLeuAlaGlnGlyAlaAlaAsnGlyTLeasn 300
890 CCACACCTTTCTGGCAGCATGATGCAAGATGCTCAGGGTGCAAGTAAATGCAATCAT 949
Qy 301 Pheuh1sgluAsnHisHisLeuHisArgAspLysSerAlaAsnLysLeuLeuAsp 320
950 TTCTACATGAAATCATCATATTCATAGAAATTAAGTGAATATCTTCTGAT 1009
Qy 321 GluAlaPheThrAlaLysLysSerAspPheGlyLeuAlaArgAlaSerGlyLysPheAla 340
1010 GAACCTTTTACTGCTAAATATCTGACTTGGCTTGACCGGCTTTCGAAAGATTTGCC 1069
Qy 341 GlnThrValMetThrSerArgLysValGlyThrThrAlaYrMetAlaProGluAlaLeu 360
1070 CAGACAGCATGATGATGCAAGAAATGTTGGGAAACAAGCTTATATGAGCAGCAAGCTTTG 1129
Qy 361 ArgGlyGluLeuThrProLysSerAspLysLeuSerPheGlyValValLeuLeuGluLeu 380
1130 CGTGAGAAATTAACACCCAAATCTGATATTTACAGCTTGGTGGTTTATCTGAATA 1189
Qy 381 11ethrthrglyLeuProAlaValAlaAspGlyHisArgGlyProGlnLeuLeuAspLys 400
1190 ATAACTGGACTTCAGCTGTGATGAGAACACCGTGAACCTCGATTATGCTGATTAATA 1249
Qy 401 GluGluLeuLeuAspGlyGluLysThrLysLeuAspTyrLysAspLysMetAsnAsp 420
1250 GAAGAATAATTGAAGATGAAGAAAGCAATGAAATTAATTAATAAAGATGAATGAT 1309
Qy 421 AlaAspSerThrSerValGluAlaMetTyrSerValAlaSerGlnCysLeuHisGlyLys 440
1310 GCTAATTCACCTTCAAGTTGAAGCTATGACTCTGTGTGCTACCAATGTCGCAAGAAAG 1369
Qy 441 LysAsnLysArgProAspLysLysValGlnGlnLeuLeuGlnGluMetThrAlaSer 460
1370 AAAAAATTAAGAGCCAGACATTAAGAAGTTTCAACAGCTGCTGCAAGAGATGACGCTTCT 1429

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: murine IL-1 receptor-associated kinase 4 (IRAK-4)
; FEATURE: CDNA
; NAME/KEY: CDS
; LOCATION: (163)..(1542)
; OTHER INFORMATION: murine IRAK-4
US-09-759-595-4
Alignment Scores:
Pred. No.: 2,026-217 Length: 1542
Score: 2017.00 Matches: 385
Percent Similarity: 90.85% Conservative: 32
Best Local Similarity: 83.88% Mismatches: 42
Query Match: 84.53% Indels: 0
Gaps: 0
US-10-001-254-16 (1-460) x US-09-759-595-4 (1-1542)
Qy 1 MetAsnLysProIleThrProSerThrTyrValArgCysLeuAsnValGlyLeuIleArg 20
163 ATGAACAAGCCGTTGACACCATGACATACATACGCAACCTTAATGTGGAGATCTTAGC 222
Db 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrLysLysLeuAlaValAlaLys 40
223 AAGCTGCGGATTTTATGATTCCTCAAGAGGGTGGAAAGAAATTAAGCAGTAGCTATCAA 282
Qy 41 LysProSerGlyLysAspAspArgTyrAsnGlnPheHisLysLeuArgPheGlyAlaLeu 60
283 AAGCGGTCGCGCAGCAGACATCAATCACTTCATTAAGAGATTGCAAGGCTTACTT 342
Qy 61 GlnThrGlyLysSerProThrSerGlyLeuLeuPheAspTyrGlyThrThrAsnCysThr 80
343 CAGACCGGAAAGAGCCCACTGTGAACTCTGTTGACTGGGGCACCAAGAACTGCACA 402
Qy 81 ValGlyAspLeuValAspLeuLeuLysGlnAsnGlyPhePheAlaProAlaSerLeuLeu 100
403 GTTGGGCACTTGTGATGATCTAGTGCACAGATTAAGAGCTGTTGCCCGCCACTCTCTG 462
Qy 101 LeuProAspAlaValProLysThrAlaAsnThrLeuProSerLysGluAlaIleThrVal 120
463 CTGCGGATCCCGTTCGCCAAACCGTCAAAAGCTGCTCTCTGAGAGAGCGGCAACAGTG 522
Qy 121 GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140
523 GCACAAACACACAGGCGCTTGTCAAGAAAGACAGACATCCGTAATGCCATATGCCGAG 582
Qy 141 LeuGlnGlnSerTyrMetProProAspSerSerSerProGluAsnLysSerLeuGluVal 160
583 CTGAACACACAGCTGCGGACGACCGGACTCTCAAGCCAGACAAAGAGTGTAGATGCC 642
Qy 161 SerAspThrArgPheHisSerPheSerPheTyrGluLeuLysAsnValThrAsnAsnPhe 180
643 AGGACACCTCGGTTCCACAGCTTCTGTTCCATGAATCTGAAGAGCATCAACAAACTTCT 702
Qy 181 AspGluArgProIleSerValGlyYasnLysMetGlyGluGlyGlyPheGlyValVal 200
703 GACGAGAACCCCGCTGCGGCTGGCAACGGATGGAGAGGGAGATTGGAGTGGTG 762
Qy 201 TyrLysGlyYrValAsnAsnThrThrValAlaValLysLysLeuAlaAlaMetValAsp 220
763 TACAAGGCGTGTGAAACAACATCGTGGCGGTGAAGAGCTCGGAGGAGATGTTGA 882
Qy 221 11ethrthrglugluLeuLysGlnGlnPheAspGlnGluLysValMetAlaLysCys 240
823 ATCAGTACTGAAAGAACTTAAGCAACAGTTTGAATCAAGAAATTAATGCAAGCGTGT 882
Qy 241 Glnh1sgluAsnLeuValGluLeuLeuGlyPheSerSerAspGlyAspAspLeuCysLeu 260

DB 883 CACACAGAGACCTGGAGAGCTGCTGCTTCCAGCAGCAGCAACCTGTGCTTA 942
QY 261 ValTyrValTyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThr 280
DB 943 GTGTATGCTTACATGCGCAACCGGTCCTGCTGACAGATCTGCTGAGATGTGATCA 1002
QY 281 ProProLeuSerTrpHisMetArgCysValValLeuAlaGlnGlyValAlaAsnGlyLeuAsn 300
DB 1003 CCACCGCTTTCCTGACACAGAGTGCAGAGTCTCAGGGGACAGCAATGCGATCAGG 1062
QY 301 PheLeuHisGlyLeuAsnHisIleHisArgAspIleLeuSerAlaAsnIleLeuLeuAsp 320
DB 1063 TTTCTGATGAATATCATCATTCATCATGATATTAAAGTCGAATATCTACTAGAC 1122
QY 321 GluAlaPheThrAlaIleValIleSerAspPheGlyLeuAlaArgAlaSerGlyLysPheAla 340
DB 1123 AAGAGCTTAACTCCCAAAATATCTGACTTTGGCTTGACAGGCTTCGCGAAGCTTAGCG 1182
QY 341 GluThrValMetThrSerArgIleValGlyThrThrAlaTyrMetAlaProGluAlaLeu 360
DB 1183 CAGACGCTCATGACCAAGCCCAATCGTGGCACAACGCTTACATGCAACCCGAAGCTTTG 1242
QY 361 ArgGlyGluIleThrProLysSerAspIleTyrSerPheGlyValValLeuLeuGluIle 380
DB 1243 CGGGAGAAATTAACCCCAATCTGACATCTACAGCTTCGCGTGGTTCTGTGGAGCTG 1302
QY 381 IleThrGlyLeuProAlaValAlaArgIleHisArgGluProGlnIleLeuLeuAspIleLys 400
DB 1303 ATTAACCGGCTGCGCGCTGTGATGAAACCGTGAACCTCACTACCTGCTGATATTAA 1362
QY 401 GluGluIleGluAspGluGluLysThrIleGluAspTyrIleAspLysLysMetAsnAsp 420
DB 1363 GAAGAGATTGAAGTAGAAGAGACATTTGAAGATTACACGATGGAAGATGAGGAT 1422
QY 421 AlaAspSerThrSerValGluAlaMetTyrSerValAlaSerGlnCysLeuHisGlyLys 440
DB 1423 GCGGACCTGCTTCGCGGAGAGCAATGATCTGCTGCTGCAAGTGTCTGATGAGAG 1482
QY 441 LysAsnLysArgProAspIleLysLysValGlnGlnLeuLeuGlnGlnMetThrAla 459
DB 1483 AAAAACAAGACGCGCAGCATTTGCAAGGTTCACAGCTGTACAAAGATGTGCT 1539
RESULT 6
US-09-764-868-249/c
; Sequence 249, Application US/09764868
; Patent No. US2002016871A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; PRIOR FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 249
; LENGTH: 1493
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-249
Alignment Scores:
Pred. No.: 1,866-99 Length: 1493
Score: 972.00 Matches: 190
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 40.74% Indels: 0
DB: 10 Gaps: 0
US-10-001-254-16 (1-460) x US-09-764-868-249 (1-1493)
QY 271 LeuAspArgLeuSerCysLeuAspGlyThrProProLeuSerTrpHisMetArgCysLys 290

DB 1491 CTAGACAGACTCTCTTGTGATGATGCTACTCACACATTTCTTGACATGAGATGCAAG 1432
QY 291 IleAlaGlnGlyValAlaAsnGlyLysPheLeuHisGlyAsnHisIleHisArg 310
DB 1431 ATTGCTCAGAGGAGCAGCTAATGCAATTTCTTCAAGAAATCATATATTCATACA 1372
QY 311 AspIleLysSerAlaAsnIleLeuLeuAspGluAlaPheThrAlaLysIleSerAspPhe 330
DB 1371 GATATTAAAGTGCAAATATCTTACTGATGAAGCTTTACTGCTAAATATCTGACTTT 1312
QY 331 GlyLeuAlaArgAlaSerGlnLysPheAlaGlnThrValMetThrSerArgIleValGly 350
DB 1311 GGCCTTGACAGGCTTCTGAGAAAGTTGGCCACAGATCATGACGAAATTTGGGCA 1252
QY 351 ThrThrAlaTyrMetAlaProGluAlaLeuArgGlyGluIleThrProLysSerAspIle 370
DB 1251 ACAACGCTTAATATGACACCAAGCTTGGCGGAGAAATACACCCAAATCTGATATT 1192
QY 371 TyrSerPheGlyValValLeuLeuGluIleIleThrGlyLeuProAlaValAspGluHis 390
DB 1191 TACAGCTTGGTGTGCTTTTACTAGAAATATTAAGCTTCCAGCTGTGATGACAC 1132
QY 391 ArgGluProGlnLeuLeuAspIleLysGluGluIleGluAspGluGluLysThrIle 410
DB 1131 CGTGAACCTCAGTTATTGCTAATATTAAAGAAATGAAGATGAAGAAAGCAATT 1072
QY 411 GluAspTyrIleAspLysLysMetAsnAspAlaAspSerThrSerValGluAlaMetTyr 430
DB 1071 GAAGATTATTAATGAATAAAGATGAATGCTGATTCACATTCAGTTCAAGATATGAC 1012
QY 431 SerValAlaSerGlnCysLeuHisGlyLysAsnLysArgProAspIleLysVal 450
DB 1011 TCTGTTGCTAGTCAATGCTGTGATGAAAAGAAATTAAGACCAAGCATTAAGAGATT 952
QY 451 GlnGlnLeuLeuGlnGlnMetThrAlaSer 460
DB 951 CAACAGCTGCTCAAGAGATGACAGCTTCT 922
RESULT 7
US-09-745-763-109
; Sequence 109, Application US/09745763
; Patent No. US20020065394A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John M.
; Lavallie, Edward R.
; Collins-Racie, Lisa A.
; Evans, Cheryl
; Merberg, David
; Treacy, Maurice
; Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 219
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genentech Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/745,763
; FILING DATE: 18-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.

REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 483 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 109:
US-09-745-763-109

Alignment Scores:

Pred. No.:	2,096-78	Length:	483
Score:	779.00	Matches:	159
Percent Similarity:	98.76%	Conservative:	0
Best Local Similarity:	98.76%	Mismatches:	1
Query Match:	32.65%	Indels:	2
DB:	9	Gaps:	0

US-10-001-254-16 (1-460) x US-09-745-763-109 (1-483)

QY 270 LeuleuaspargleuseSerCyaleuaspGlyThrProProleuSerTriphMetarGys 289
DB 3 TTGCTAGACACACTCTCTTGGTGGATGGTACTCCACACCTTCTGGACATGAGATGC 62
QY 290 LysllealaglinalaalaasnGlylleasnphleuHsgluasnHihisIlehis 309
DB 63 AAGATTGCTCAGGGGAGCTAATGGCATCAATTTTCAACATGAATAATCATCATTCAT 122
QY 310 ArgaspilleuSerleuAasnIleleuaspGluualaphethrAlaIleSerasp 329
DB 123 AGAGATATTAAGCAAAATATCTTACTGATGAAGCTTTTACTGCTAAATATCTGAC 182
QY 330 PheglleuAlaArgAlaSerGluys-PheAlaGlnThrValMetThrSerAglIleVa 349
DB 183 TTGGCCCTGCACGGGCTTGCAGAAAGTTTGCCACAGCATCATGACAGAAATTTGT 242
QY 349 lGlyThrThrAlaTyrrMetAlaProGluAlaIleuArgGlyGluIleThrProlysSerAs 369
DB 243 GGGAAACAAGCTATATGCGACCAAGAGCTTTGGTGAGAAATTAACACCAATCTGA 302
QY 369 PileTySerPheGlyValIleleuGluIleIleThrGlyLeuProAlaValaspG 389
DB 303 TATTTACAGCTTTGGTGGTGTGTTTACAGAATATATACCTGACCTCGATGATGA 362
QY 389 uHieArgGluProGlnleuLeuAspIleIleGluGluIleGluaspGluGluIleSTH 409
DB 363 ACACCGTGAACCTCAGTTTATGCTAGATATTTAAGAGAAATGAGATGAAGAAAGAC 422
QY 409 rIleGluaspTyrlleaspIlySmetAsnaspAlaaspSerThrSerValGluIle 429
DB 423 ATT-GAAGATATATTTGATTAAGAAAGATGATGATGATGATTCACCTTCACTGAAGCAT 481
QY 429 c 429
DB 482 G 482

RESULT 8

US-09-833-790-149
; Sequence 149, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tonglong
; APPLICANT: Secrist, Heather
; APPLICANT: Mohamath, Radoch
; APPLICANT: Indrias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.512
CURRENT APPLICATION NUMBER: US/09/833,790
CURRENT FILING DATE: 2001-04-11
NUMBER OF SEQ ID NOS: 440
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 149
LENGTH: 501
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(501)
OTHER INFORMATION: n = A,T,C or G
US-09-833-790-149

Alignment Scores:

Pred. No.:	1,916-68	Length:	501
Score:	691.00	Matches:	151
Percent Similarity:	94.44%	Conservative:	2
Best Local Similarity:	93.21%	Mismatches:	7
Query Match:	28.96%	Indels:	5
DB:	9	Gaps:	0

US-10-001-254-16 (1-460) x US-09-833-790-149 (1-501)

QY 1 MetAsnlyspProIleThrProSerThrTyrrValArgCysleuAenValGlyleuIleArg 20
DB 21 ATGAACAACCCATTAACACCATCAATATGTCGGCTCAATGTTGGACTAATTAGG 80
QY 21 LysleuSeraspPheIleaspProGlnGlyTyrrlysuIleuAlaIleIlys 40
DB 81 AAGCTGCAGATTTTATGATCTCCAGAAAGATGAAAGAGTTAGCTGATGATTTAA 140
QY 41 LysProSerGlyaspAspArgTyrrAsnGln-PheHisIleArgThrPheGluAlaIleu 60
DB 141 AAACCATCTGCTGATGATGATGATCAATCAAGTTTCAATAGAGATTTGAAGCATT-CT 199
QY 60 uGlnThrGlySerProThrSerGluIleuPheAspTrp-GlyThrThrAsnGly 80
DB 200 TCAAACTGGAAAAAGTCCACTTCTGTAATATCTGTTGACTGGGGGACCAAAATGGA 259
QY 80 hrValGlyaspLeuValAspIleuIleGlnAenGluPhePheAlaProAlaSerIleu 100
DB 260 CAGTTGATGATCTTGTGATCTTTGATCCAAATGAAT- TTGCTCTCGCGAGCTTT 318
QY 100 euleuProAspAlaValProlySerThrAlaasnThrleuProSerlyGluAlaIleThrV 120
DB 319 TGCTCCAGATGCTGTTCCCA- ACTGCTAATACACTTCTTAAGAAAGCTATTAAG 377
QY 120 aGlnGlnlyGlnMetProPheCysAspIysAspArgThrLeuMetThrProValGlnA 140
DB 378 TTCAGCAAAAACAGATGCCCTTTCTGTGACAAAGACAGACATTCATGATCACCTGTGANA 437
QY 140 snleuGlnIleSerTyrrMetProProAspSerSerSerProGluAsnlySerleuGluV 160
DB 438 ATCTTGAACAAGCTATATGCTGACCTGATCTTCAATCCAAANAAATTAAGTTTAAG 497
QY 160 a 160
DB 498 TT 499

RESULT 9

US-09-966-451-10
; Sequence 10, Application US/0966451
; Publication No. US20030087856A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; TITLE OF INVENTION: ANTISENSE MODULATION OF IL-1 RECEPTOR-ASSOCIATED KINASE-4 EXPRES
; FILE REFERENCE: RTS-0324
; CURRENT APPLICATION NUMBER: US/09/966,451
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 88

```

; SEQ ID NO 10
; LENGTH: 31000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-966-451-10

```

Alignment Scores:

```

Pred. No.: 2,44e-54      Length: 31000
Score: 590.00           Matches: 166
Percent Similarity: 26.94%      Conservative: 4
Best Local Similarity: 26.31%      Mismatches: 6
Query Match: 24.73%             Indels: 455
DB: 11                      Gaps: 2

```

US-10-001-254-16 (1-460) x US-09-966-451-10 (1-31000)

```

QY      53  ILAAYARPhPheGluAlaLeuEulInThrGlyLysSerProThrSerGluLeuPhe 72
      12963 TTAAGGAGATTGAGCATTTACTTCAAACTGAAAAAGTCCCACTTGAAATTAAGTTT 13022
QY      73  AAPTTCGLYThrPheCysThrValGlyAspLeuValAspLeuEulIegIAsnGlu 92
      13023 GACTGGGGCACCAAAATTCACAGTTGGATCTTGATCTTTTGATCCAAATGAA 13082
QY      93  PhePheAlaProAlaSerLeuLeuLeuPro----- 102
      13083 TTTTTCCTCTCGAGTCTTTTTCCTCCAGTAACTGATGACCAAGGTGTCCACA 13142
QY      102 ----- 102
Db      13143 ATTAGGTGGAAGAACAAATGCGAGAAATATAAATGTTCTTACTCTTCTTTTTC 13202
QY      102 ----- 102
Db      13203 TCATAGTAGATGAAGCTTACATTGAGAGTCCCTTCTTCAGCATCTCAACTTTTA 13262
QY      102 ----- 102
Db      13263 AAAAGCAGCACAGACAAAGACACTGTGACTGTCTGCTAAGTGATAGAGCTTGTA 13322
QY      102 ----- 102
Db      13323 AGAGTTAGATAGTTTGTGCCAACAGAGATTTAGAGAAAGACTTCATACCTTGTGCT 13382
QY      102 ----- 102
Db      13383 TAGGCTGTAAAGAAATTAATTAATTTGAGTCTCTTTTTCATCTTCAACTTCTA 13442
QY      102 ----- 102
Db      13443 CCCTGATGGAGCTCTATATCATTAATTTTAAAAAATGTATCTGATGTAGTGTCTA 13502
QY      102 ----- 102
Db      13503 GTTGCTTCCTTAAGACTTTCTTTTTCCTTGTAGTCTTTCACACAGGTGTCTCA 13562
QY      102 ----- 102
Db      13563 ACCTTATCATGCATTAAGAAATCACTTAGGTTCTGGTTAAACATGTAAGTTTCTAGGGTCT 13622
QY      102 ----- 102
Db      13623 GTTCCAAATCATGTATATCAAGAAATTCGAAATTCGGTTTAAGTCTATATTTCTTAAACAG 13682
QY      102 ----- 102
Db      13683 GTGCTTCAGTGCCTGTATATAGGTGCTATGATCAGCTGTTGAAAGACATGCTTAC 13742
QY      102 ----- 102
Db      13743 TCTCTTACCACATCTTCATTATACTTACAGATTTCTTACTAGCAGCTTCTGTGTGTC 13802

```

```

QY      102 ----- 102
Db      13803 TGTGAGATATGAGACCAACTGTAGAAATGATATTAATTAAGAACAAAGTTTCTA 13862
QY      102 ----- 102
Db      13863 GTTTAACTTTTTCACAAACACTTTTCTTACTGAAAAACCACCTGTATCTTACTTCAATT 13922
QY      103  ----AapAlaValProLysThrAlaAsnThrLeuProSerLysGluAlaIleThrValGly 121
      13923 GTTAGATGCTGTTCCTCCAAAACAGCTATATCACTACTCTTCAAAAGACATTAACATCTCA 13982
QY      121  nGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsnLe 141
      13983 GCAAAAACAGATGCTCTTCTGTGACAAAGACAGACATGATGACACCTGTGCGAATCT 14042
QY      141  uGlnLysSerLysMetProProAspSerSerProGlnLysLeuSerLeuGluValIse 161
      14043 TGAACAAACCTATATGCGACCTGACTCTCAAGTCCAGAAATPAATPAAGTTTGAAGTTAG 14102
QY      161  rAepThrArg----- 164
      14103 TGATACACGTAAAGTAACATTTTCAGTCTTTCACACTAGGATTTGTCATTAGACTACCA 14162
QY      164 ----- 164
      14163 GTGCTTTAAAGAAAGCTCTGTGCTTTTGTGTGACAGCAATACAGGACACATGGCAA 14222
QY      164 ----- 164
      14223 TAGCTTTTGTGAGTTGTTTCTCTCGATATATTAAGAACCATCTTCAATTGATTAATC 14282
QY      164 ----- 164
Db      14283 AGTGATTAGAGCAATAGACATGCAACAAACAGAGAAAGCTATGAAAAAATAAGTAA 14342
QY      164 ----- 164
      14343 ATATTTACATACTTGAAAGGTCACTTTTAAATATAAATATATGACTAGAGTTTGGG 14402
QY      164 ----- 164
      14403 TAGTAGACTAGACCCACTCAATGTTGATCTAGAAATTTGGGTGGTGACTACAGTGT 14462
QY      164 ----- 164
      14463 CAGAGGTAGATCACCAACAGAGTTCTCAGAAAAAATCTATATTTGATTTAGAAA 14522
QY      164 ----- 164
      14523 TGGTTTATCTCTCTCATCTGTCTAATTCAAAATCATTAATATTTGCATATATGTGA 14582
QY      164 ----- 164
      14583 ATCTCAAAATGAGAAAAATATTAATTAATTAAGCAATTTTAAATTTAAGCATGTTT 14642
QY      165  -----PheHisSerPheSerPheTyrGluLeuLysAsnValThrAs 178
      14643 TTCTTATTTTGACATAGTTTTCACAGTTTTCATTTTATGATTAAGAAATGTCACAAA 14702
Db      14703 TAACTTTGATGAACGACCACTTCTGTGTGTATTAATAATGGAGAGGAGGATTTGG 14762
QY      198  yValValIlyrLysGlyTyrValAsnAsnThrThrValAlaValLysLysLeuAlaIse 218
      14763 AGTTGTATATTAAGGCTACGTAAATTAACAACATGTGGCAGTGAAGAGCTTCACACAGT 14822
QY      218  tValAspIleThrThrGluLeuLysGln 228
      14823 AAGTTATATTTTCAGGAATTAATAAGAAAGAG 14853

```

RESULT 10


```
US-10-001-254-5
; Sequence 5, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Florentino, Loredana
; APPLICANT: Lee, Sung Hyung
; APPLICANT: Roch, Wilfred
; APPLICANT: Steiner-Liewen, Frank
; TITLE OF INVENTION: No. US20030049702A1e1 Death Domain Proteins
; FILE REFERENCE: P-LJ 5037
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(294)
US-10-001-254-5

Alignment Scores:
Pred. No.: 3,77e-49 Length: 294
Score: 517.00 Matches: 98
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 21.67% Indels: 0
DB: Gaps: 0

US-10-001-254-16 (1-460) x US-10-001-254-5 (1-294)

QY 9 ThrTyrValArgCysLeuAenValGlyLeuIleArgIysLeuSerAapPheIleAapPro 28
DB 1 ACAATATGCGCTCTCAATGTTGACTATAGAAAGTGCAGATTTTATGATCT 60
QY 29 GlnGluGlyTrrpIysIysLeuAlaValAlaIleIysIysProSerGlyAaspAapArgTyr 48
DB 61 CAAGAAGATGGAAGAGTGTAGCTGATTAATAAAACCATCTGTGATGATGATGATAC 120
QY 49 AsnGlnPheHisIleArgArgPheGlnAlaLeuLeuGlnThrGlyIysSerProThrSer 68
DB 121 AATCAGTTTCACTAAGAGATTTGAAAGCATTTCTCAACTGAAAGTCCCACTTCT 180
QY 69 GluLeuLeuPheAapTrrpGlyThrThrAsnCysThrValGlyAapLeuValAapLeuLeu 88
DB 181 GAATTACTGTTGACTGCGGACCAACAATTCACAGTGGATGATCTTGTGATCTTTTG 240
QY 89 IleGlnAenGlnPhePheAlaProAlaSerLeuLeuLeuProAapAlaValPro 106
DB 241 ATCCAAATGATTTTGTCTCCGCAAGTCTTTGTCTCCAGATGCTGTTCC 294

RESULT 11
US-09-880-107-2222
; Sequence 2222, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880.107
; CURRENT FILING DATE: 2001-06-14
```

```
PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2222
; LENGTH: 3590
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L76191
US-09-880-107-2222

Alignment Scores:
Pred. No.: 2.43e-47 Length: 3590
Score: 515.50 Matches: 156
Percent Similarity: 46.14% Conservative: 89
Best Local Similarity: 29.38% Mismatches: 192
Query Match: 21.61% Indels: 94
DB: Gaps: 19

US-10-001-254-16 (1-460) x US-09-880-107-2222 (1-3590)

QY 4 ProIleThrProSerThrTyrValArgCysLeuAenValGlyLeuIleArgIysLeuSer 23
DB 115 CCCCAGGCGCCAGACACTTCTTGTACAGTCCGCCCTGATGATGATGATGATGATGATGAT 174
QY 24 AapPheIleAapProGlnGlnGlyTrrpIysIysLeuAlaValAlaIleIysIysProSer 43
DB 175 AGT-----GATGACGCGCCCTGAGCCGCGGACGCTG-----GTGCAAGTCCGCCCT 222
QY 44 GlyAaspAapArg-TyrAsnGlnPheHisIleArgArgPheGlnAlaLeuLeuGlnThrGly 63
DB 223 -----GATGTGCGCGACCAACGACGCTGCGG-----GTGCGACGCGCTC 264
QY 63 YLysSerProThrSerGlnLeuLeuPheAapTrrpGlyThrThrAsnCysThrValGlyAs 83
DB 265 CGGACGCGCGACCGCGCGCGCTGCTGCGCTGATCAACCGACCGCGCTGCGCGCA 324
QY 83 PLeuValAapLeuLeu-----IleGlnAenGlnPhe 93
DB 325 CTTGCGCAATCTCTCAACGACCTGAGCTCCGTCGCGCGGACATCATCAAGCTG 384
QY 93 ePheAlaProAlaSerLeuLeuLeuProAapAlaValProIysThrAlaAsnThrLeuP 113
DB 385 GCACCTCCCGCGCGCGCTTCCGTCGCCGACCACTGCGCGCGCGCGCGCGCATCC 444
QY 113 oSer---LysGlnAlaIleThrValGlnGlnIysGlnMetProPheCysAapIysAapAr 132
DB 445 TGCACCGCGCGCGCGCGCGCGCTGAGACCGCGCGGAGTTGCCA-----TCTCAGCCTC 498
QY 132 gThrLeuMetThrPro-----LeuGlnValSerAapThrArgPheHisSerPh 168
DB 499 CACTTCTCTCTCCAGGCTTTTCCAGGCTCCCAACCATTCAGGCGCTGAGCTCGGCT 558
QY 138 -ValGlnAenGlnGlnSerTyrMetProAapSerSerSerProGlnIysIys 157
DB 559 GGTTCAGAGCCCTGCTTCCCTGCGCTCAACCGCATCTCCACCTTCTTCTTCA 618
QY 157 r-----LeuGlnValSerAapThrArgPheHisSerPh 168
DB 619 GCCAGCGCCAGAGAGCTCACTGCTCTCTGCGAGGAGCGCGCGCTTCCGTTTGGCTG 678
QY 168 eSerPheTrrpGlnLeuIysAenValThrAsnAenPheAapGlnArgProIleSerValG 188
DB 679 GCCCTCTGAGATTTCCCGGCGCACCAACTTCTCGAGAGAG----- 724
QY 188 yGlnAenIysMetGlyGlnGlyGlyPheGlyValValIlyTrpIysGlyTyrValAAsnA 208
DB 725 ----CTCAAGATCGGGAGGAGGCTTGTGGTGGCGGCGGCGGTGAGGAAC 780
QY 208 rThrValAlaValIlyLysLeuAlaIleMetValAapIleThrThrGlnGlnLeuIysG 228
```

```

Db      781 GGTATATGCTGTAGAGGCTGAAGAGAAAGCGTGAACCTGAGTGAAGCTGACAGTGAAGCA 840
Qy      228 nglnphearpglnlulleyvalmetalaalyscsyglinhiegluasnleuvalglule 248
Db      841 GAGCTTCCTGACCGAGGTGAGCAGCTTCAGGTTTCCTGACCCAAACATTGTGAGCTT 900
Qy      248 uleuglypheserSerAsp1yAspAspLeuCySleuValTyValTymerProAsn1 268
Db      901 TGTGTGCTACTGTGCTCAGAACGGCTTTACTGCTGCTGTGTATAGGCTTCTGCTCCAAAG 960
Qy      268 ySerleuLeuAspArgLeuSerCys---LeuAspGlyThrProProLeuSerTrpHis 287
Db      961 CTCCTCGAGAGACCGTCTCCACTGACAGACCCAGCCAGCCGCTCCCTCTGAGCCCTCA 1020
Qy      287 TAGGCSySyllealaglmglyalalaanglylleasnPhelenuhieglnsnhish 307
Db      1021 GCGACTGACATCTTCTGAGTACAGCCCGGCAATTCACTTCTCATCAGAGACAGCC 1080
Qy      307 s-----lleHsArgAsp1lelySerAlaAsn1leuLeuAspGlyAlaPheThrAl 325
Db      1081 CAGCTCATCTCAGAGACATCAAGAGTTCCAACTCTTGATGAGAGCTGACACC 1140
Qy      325 alySileSerAspPheGlyLeuAlaArgAlaSerGlyLysPheAla----- 340
Db      1141 CAGCTGGAGACTTGTGCTGCGCCGTTCCAGC---CGCTTGGCGGCTCCAGCCAG 1197
Qy      341 -GlnThrValMetThrSerArg-----lleValGlyThrThrAlaTyMetAlaPr 357
Db      1198 CCGAGCAGCATGATGATGCGCCGACACAGACAGATGCGGGGCACTGCTGACTGCTG 1257
Qy      357 ogliuAlaLeuArg---GlyGluIleThrProLysSerAsp1leTySerPheGlyValVa 376
Db      1258 GGGATCATATCAAGACGGGAAGGCTGCTGTGACACGGACACCTTCAGCTTTGGGGTGT 1317
Qy      376 lleuLeuGluIleIleThrGlyLeuProAlaValAspGluHis---ArgGluProGlu 395
Db      1318 AGTGTAGAGACCTTGGCTGTGACAGGCTGTGTAAGACGACGATGTCAGACCAAGTA 1377
Qy      395 uleuLeuAsp1lelySgluIleGluAspGlu-----G1 407
Db      1378 TCTGAAGAAGCTGCTGTGAGAGAGAGCTGAGAGGCTGAGTGGCTTTGAGAAGACCCA 1437
Qy      407 ulysThrIleGlu----- 411
Db      1438 GAGCAGACTCAGACAGCTGTGCTGACATGCTGGCTGCCATGCCATGACGAT 1497
Qy      412 -----AspTyrlleAspLysLysMetAsnAspAlaAspSerThrSerValGluAla 429
Db      1498 CTACAGAAGACACTGAGACCCAGGCCCGGCTGCCCTGACCTGAGCTGGGGCT 1557
Qy      429 tTySerValAlaSerGlnCySleuHisGlyLysLysAsnLysArgProAsp1lelySyl 449
Db      1558 GGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1617
Qy      449 sValGlnGlnLeuLeuGlnGluMetThrAla 459
Db      1618 GGTGTACGAGAGGCTAGAGAACTGACAGCA 1648

```

```

; LENGTH: 3590
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (80)...(2218)
; US-10-167-034-4

Alignment Scores:
Pred. No.: 2,43e-47 Length: 3590
Score: 515.50 Matches: 156
Percent Similarity: 46.14% Conservative: 89
Best Local Similarity: 29.38% Mismatches: 192
Query Match: 21.61% Indels: 94
DB: 12 Gaps: 19

US-10-001-254-16 (1-460) x US-10-167-034-4 (1-3590)
Qy      4 ProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArgLysLeuSer 23
Db      115 CCGCGCGCCGACGACTTGTATGAGGTGCGCGCTGCTGATGATGCTGCTGCTGCTGCTG 174
Qy      24 AspPheIleAspProGlnGluGlyTyTrpLysLysLeuAlaValAlaIleLysLysProSer 43
Db      175 AGT-----GATGACGCGCTGAGAGCCGCGCACTG---GTGCCAGTTGCGCGCTCT--- 222
Qy      44 GlyAspAspArg-TyrAsnGlnPheHisIleArgArgPheGluAlaLeuGlnThrG1 63
Db      223 -----GATGTCGCGGACACGACGAGCTGCGG-----CTGTGCGAGCGCTC 264
Qy      63 LysSerProThrSerGlnLeuLeuPheAspTrpGlyThrThrAsnCyThrValGlyLys 83
Db      265 CCGGACGCGCCAGCGCCAGCGCTCTGTGCGCTGTGATCAACCGGACCGCCGCTGTGCGCA 324
Qy      83 pleuValAspLeuLeu-----lleGlnAsnGluPh 93
Db      325 CCTCGTGACATCTCAGCAGCAGCTGCTGCTGCGGAGCATCATGACAGCTG 384
Qy      93 ePheAlaProAlaSerLeuLeuLeuProAspAlaValProLysThrAlaAsnThrLeuPr 113
Db      385 GCACTCTCCGCGCGCTTCCCTGCTCCGCGACACCACTGCCGAGGCCAGCGACGATCCC 444
Qy      113 oSer---LysGluAlaIleThrValGlnGlnLysGlnMetProPheCyAspLysAspAr 132
Db      445 TGCACCGCGGAGCGCGAGCGCTGAGAGCCCGGAAAGTTGCCA-----TCTCAGACTTC 498
Qy      132 gThrLeuMetThrPro----- 137
Db      499 CACTTCTCTCTCCCGCACTTTTCCAGGCTCCAGACCCATTGAGGCTGAGCTCGGCT 558
Qy      138 -ValGlnAsnLeuGlnGlnSerTymerProProAspSerSerSerProGluLysLys 157
Db      559 GGTTCAGACCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
Qy      157 r-----leuGluValSerAspThrArgPheHisSerPh 168
Db      619 GCGAGGCCAGAGAGCTCAGTGTCTCTGACAGGAGCCCGCCCTTCCGTTTGCTG 678
Qy      168 eSerPheTyrgluLeuLysAsnValThrAsnAsnPhaAspGluArgProIleSerValG1 188
Db      679 GCGCTCTGTGAGATTCCCGGGGACCCACCAACTTCTGAGAGAG----- 724
Qy      188 yGlyAsnLysMetGlyGlnGlyGlyPheGlyValValTyTrpLysGlyTyValAsnAsn 208
Db      725 -----CTCAAGATGCGGAGAGGTGGCTTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 780
Qy      208 rThrValAlaValLysLysLeuAlaIleMetValAspIleThrThrGlnGluLeuLysG1 228
Db      781 GGTGTATGCTGTAGAGGCTGAGAGAGACGTCGATGCTGAGTGAAGCTGACAGCA 840
Qy      228 nglnphearpglnlulleyvalmetalaalyscsyglinhiegluasnleuvalglule 248

```

```

Db      841 GAGCTTCCTGACCGAGGTGAGCAGCTGTCCAGGTTTCGTCACCAACATTGTGACTT 900
Qy      248 uleuGlypheSerSerAspGlyAspAspLeuCySerValTyrValTyrMetProAnGI 268
Db      901 TGCTGGCTACTGTGCTCAAGAACGGCTTACTGCTGGGTGATGAGGCTTCCGCCAACGG 960
Qy      268 ySerLeuLeuAspArgLeuSerCys---LeuAspGlyThrProProLeuSerTrpHisMe 287
Db      961 CTCCTGGAGGACCGTCTCCACTGCCAGACCCAGCCCTGCCACTCTCTCTGCGCTCA 1020
Qy      287 tArgCysLysIleAlaGlnGlyAlaAlaAsnGlyIleAsnPhelLeuHisGluAsnHis 307
Db      1021 GCGATGTGACATCTCTTCTGGGTACAGCCGGGCAATTGATTTCTACATCAGACAGCCC 1080
Qy      307 s-----lleHsArgAspIleLysSerAlaAsnIleLeuLeuAspGlyAlaAlaPheThrAl 325
Db      1081 CAGCTTCATCCATGAGACATCAAGATTCACACCTCTTGATGAGAGGCTGACACC 1140
Qy      325 alysIleSerAspPheGlyLeuAlaArgAlaSerGlyLysPheAla----- 340
Db      1141 CAAGCTGGAGACTTTGGCTGCGCCGGTTCAGC---CGCTTGGCGGGTCCAGCCCGAG 1197
Qy      341 -GlnThrValMetThrSerArg-----lleValGlyThrThrAlaTyrMetAlaPr 357
Db      1198 CCAGAGCAGCATGGTGGCGCCGACACAGACAGATGGCGGGCACCTGGGCTTACCTGCCGA 1257
Qy      357 oGluAlaLeuArg---GlyGluIleThrProLysSerAspIleTyrSerPheGlyValVa 376
Db      1258 GGAGTACTCAAGACGGGAGAGGCTGGCTGTGAGACCGGACACTTCACCTTTGGGGTGT 1317
Qy      376 lleuLeuGluIleIleThrGlyLeuProAlaValAspGluHis---ArgGluProGlnLe 395
Db      1318 AGTCTGAGAGACCTTGGCTGTGAGAGGCTGTGAAAGCGACAGGTGCAGAGCAACAGTA 1377
Qy      395 uLeuLeuAspIleLysGlnGluIleGluAspGlu-----G1 407
Db      1378 TCTAAGAGCCTGTGTGAAAGAGAGGCTGAGAGGCTGAGTGTGCTTTGAGAACACCCA 1437
Qy      407 uLysThrIleGlu----- 411
Db      1438 GAGCACTGCAGACAGGTCTGGCTGCAGATGCTGGGCTGCTCCATCGCCATGCAAT 1497
Qy      412 -----AspTyrIleAspLysLysMetAsnAspAlaAspSerThrSerValGluAlaMe 429
Db      1498 CTACAAAGACACCTGGAACCCAGGCGCCGCTGCCACCTGAGCTGGGCTGGGCTT 1557
Qy      429 tTyrSerValAlaSerGlnCysLeuHisGluLysLysAsnLysArgProAspIleLys 449
Db      1558 GGGCCAGCTGGCTGTGCTGTGCTGCACCGCGGGCCAAAGAGAGGCTCTCTATGACCCA 1617
Qy      449 sValGlnGlnLeuLeuGlnGluMetThrAla 459
Db      1618 GGTGTACGAGAGGCTAGAGAGCTGCAGGCA 1648

```

```

RESULT 13
US-10-101-510-88
: Sequence 88, Application US/10101510
: Publication No. US20030148295A1
: GENERAL INFORMATION:
: APPLICANT: WANG, YIXIN
: TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
: FILE REFERENCE: 15117.0012
: CURRENT APPLICATION NUMBER: US/10/101.510
: PRIORITY FILING DATE: 2002-03-20
: PRIOR APPLICATION NUMBER: 60/276,947
: NUMBER OF SEQ ID NOS: 805
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 88
: LENGTH: 3590
: TYPE: DNA
: ORGANISM: Homo sapiens

```

```

US-10-101-510-88
Alignment Scores:
Pred. No.: 2,43e-47 Length: 3590
Score: 515.50 Matches: 156
Percent Similarity: 46.14% Conservative: 89
Best Local Similarity: 29.38% Mismatches: 192
Query Match: 21.61% Indels: 94
DB: 13 Gaps: 19

US-10-001-254-16 (1-460) x US-10-101-510-88 (1-3590)
Qy      4 ProIleThrProSerThrTyrValArgCysLeuAsnValGlyLeuIleArgLysLeuSer 23
Db      115 CCGCGGCGCCAGACACTTCTTGTACAGGTGCGCCCTGGGTATGATGCGCTTTCACAA 174
Qy      24 AspPheIleAspProGlnGlnGlyTyrLysLeuAlaValAlaIleLysProSer 43
Db      175 AGT-----GATGAGCGCCCTGGAGCGCCGCACTG---GTGCAATTCGCCGCCCT 222
Qy      44 GlyAspAspArg-TyrAsnGlnPheHisIleArgArgPheGluAlaLeuLeuGlnThrG1 63
Db      223 -----GATGTCGCGCAGACAGACCGAGCTGCGG-----CTGTGCGAGCGCTC 264
Qy      63 YLysSerProThrSerGluLeuLeuPheAspTyrGlyThrThrAsnCysThrValG1As 83
Db      265 CGGAGCAGCGACGAGCGAGCTGTGGCCCTGATCAACCGACGCCGCTGTGGCCGA 324
Qy      83 pLeuValAspLeuLeu-----lleGlnAsnGluPh 93
Db      325 CTTGTGTCATCATCTTCACGACCGACGCTGCTCCGCGGGGAAATCATCAACAGCTG 384
Qy      93 ePheAlaProAlaSerLeuLeuLeuProAspAlaValProLysThrAlaAsnThrLeuPr 113
Db      385 GCACCTTCCCGCCCGCTTCCGTCCCGCAGCACACACTGCCCGAGGCCAGCAGCATCCC 444
Qy      113 oSer---LysGluAlaIleThrValGlnGlnLysGlnMetProPheCysAspLysAspAr 132
Db      445 TGCAACCGCCGAGGCGAGGCTGTGAGCCCCCGAAGTTCCA-----TCTCAGCTTC 498
Qy      132 gThrLeuMetThrPro----- 137
Db      499 CACCTTCTCTCTCCCAAGCTTTTCCAGGCTCCAGACCCATTCAGAGGCTGAGCTCGGCT 558
Qy      138 -ValGlnAsnLeuGlnGlnSerTyrMetProProAspSerSerProGluAsnLysSe 157
Db      559 GGTTCAGAGCCCTGCTTCCCTGTGAGCTCCACCGCATTCACGCCCTTCTTCCAA 618
Qy      157 r-----LeuGluValSerAspThrArgPheHisSerPh 168
Db      619 GCCAGGCCAGAGAGCTCACTGCTCTCTGCAAGGAGCCCGCCCTTCCGTTTGGCTG 678
Qy      168 eSerPheTyrGluLeuLysAsnValThrAsnAsnPheAspGluArgProIleSerValG1 188
Db      679 GCCCTCTGTGAGATTTCCCGGGGCAACCAACTTCTGAGAGAG----- 724
Qy      188 yGlyAsnLysMetGlyGlnGlyGlyPheGlyValValTyrLysGlyTyrValAsnAsnTh 208
Db      725 ----CTCAAGATGGGGAGGGGTGGCTTGGGTGCGGTACCGGGCGGTGATGAGAAAC 760
Qy      208 tThrValAlaValLysLysLeuAlaAlaMetValAspIleThrThrGluLeuLysG1 228
Db      781 GGTGTATGCTGTGTAAGAGCTGAAGAGAAAGCTGACCTGAGATGACCTGCACTGAAG 840
Qy      228 ngInPheAspGlnGluIleLysValMetAlaLysCysGlnHisGluAsnLeuValG1Gue 248
Db      841 GAGCTTCGAGACCGAGGTGAGAGAGCTGTCCAGGTTTCGTCACCAACATTGTGACTT 900
Qy      248 uleuGlypheSerSerAspGlyAspAspLeuCySerValTyrValTyrMetProAnGI 268
Db      901 TGCTGGCTACTGTGCTCAAGAACGGCTTACTGCTGGGTGATGAGGCTTCCGCCAACGG 960
Qy      268 ySerLeuLeuAspArgLeuSerCys---LeuAspGlyThrProProLeuSerTrpHisMe 287

```

```

Db      961 CTCCTCGAGGAGGACGCTCTCCACATGCAAGACCCAGGCTGCGCCACCTCTCTCGGCTCA 1020
Qy      287 tArgCySlySllleAlnglYalAlaenglylAlaenPheleuHiSgluamhiShi 307
Db      1021 GGCACCTGACATCTCTCTGCTGACAGCCCGGCAATTCAGTTTCTACATCAGGACAGCC 1080
Qy      307 s-----lleHlaArgAspIleYSerAlaenileleuAAspGluAlaPheThraI 325
Db      1081 CAGCCTCATTCATGAGACATCAAGATTCAACGCTCTTCTGATGAGAGCTGACACC 1140
Qy      325 alySllSerAspPheGlyLeuAlaArgAlaSerGluYsPheAla----- 340
Db      1141 CAGCTGGAGACTTTGGCTGCGCCGCTTACG---CGCTTGGCGGGTCCAGCCCGCA 1197
Qy      341 -GlnThrValMetThSerArg-----lleValGlyThrThraIAlaYrMetAlaPr 357
Db      1198 CCGAGACAGATGTGTGGCCCGGACACAGACAGTGGCGGACCCCTGACCTGCGCCA 1257
Qy      357 ogIuaAlaLeuArg---GlyGluIleThrProlySerAspIleYrSerPheGlyValVa 376
Db      1258 GAGATCATCAACACGGAGAGCTGCTGACACAGACACCTTCACTTTGGGTGT 1317
Qy      376 lleuLeuGluIleIleThrGlyLeuProAlaValAspGluHi---ArgGluProGluLe 395
Db      1318 AGGTCTAGAGACTTGGCTGTGACAGGCTGTGAAGAGCAGCAGTTCAGCAAGTA 1377
Qy      395 uleuLeuAspIleGlyGluIleGluAspGlu-----Glu 407
Db      1378 TCTGAAGAAGCTGTGTGAAGAGAGGCTGAGAGGCTGAGTGGCTTTGAGAAGACCCA 1437
Qy      407 uLySThrIleGlu----- 411
Db      1438 GAGCACTCTCAAGCAGAGTGTGCTGCATGCTGCGCTGCCATGCCCATGCAAT 1497
Qy      412 -----AspTyrlleAspIleYsMetAlaAspSerThSerValGluAlaWe 429
Db      1498 CTACAGAGACACTGTGACCCAGCCCGGCTGCTGCCACCTGAGCTGGGCTCT 1557
Qy      429 tYrSerValAlaSerGlnYsleuHiSgluYslyAsnlyAsnlyAsnlyAsnly 449
Db      1558 GGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1617
Qy      449 sValGlnGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 459
Db      1618 GGTGTACGAGAGGCTGAGAGAGCTGACAGCA 1648

RESULT 14
US-09-873-367C-1003
; Sequence 1003, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Andrews, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Edner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0

```

```

; SEQ ID NO 1003
; LENGTH: 3590
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-1003

Alignment Scores:
Pred. No.: 2,436-47
Score: 515.50
Percent Similarity: 46.14%
Best Local Similarity: 29.38%
Query Match: 21.61%
DB: 13 Gaps: 19

US-10-001-254-16 (1-460) x US-09-873-367C-1003 (1-3590)

Qy      4 ProIleThrProSerThrTyValArgCySleuAenValGlyLeuIleArgIysleuSer 23
Db      115 CCCCAGCCGCGCAGCACTTCTGTACAGAGTGGCCCTGGGTCATGTGCGCTTCA 174
Qy      24 AspPheIleAspProGlnGluGlyTrpIyslySleuAlaValAlaIleYslySproSer 43
Db      175 AGT-----GATGACGCTCTGAGAGCCGCGCACTG---GTCCAGTTGCGGCT 222
Qy      44 GlyAspAspArg-TyrAsnGlnPheHisIleArgArgPheGluAlaLeuGlnThrG 63
Db      223 -----GATCGTGGCGGACCAAGCCAGCTGCGG-----CTGTGAGAGGCTC 264
Qy      63 YlySerProThrSerGlnLeuLeuPheAspTrpGlyThrThrAnCySthValGlyS 83
Db      265 CCGGACGCGCAGCGGCGAGGCTGCTGTGCGCTGTGATCAACCGCAACGCGGTGGCCGA 324
Qy      83 PleuValAspLeu-----lleGlnAsnGluPh 93
Db      325 CTTGTGACATCTTCAACCACTGCACTGCTGCTGCGGAGACATATCAACGCTG 384
Qy      93 ePheAlaProAlaSerIleuLeuProAlaValProlySthAlaAsnThrIeudr 113
Db      385 GCAACCTCCCGCCGCTTCCGTCCCGCAGCACACTGCCCGGAGCCAGCAGCATCC 444
Qy      113 oSer---lySgluAlaIleThrValGlnGlnlySgluMetProPheCyAspIlyAspAr 132
Db      445 TGCACCCGCGAGGCGGAGGCTGAGACCCCGGAAGTTGCCA-----TCCCTAGCCTTC 498
Qy      132 GThrIleuMetThrPro----- 137
Db      499 CACTTCTCTCTCCCAAGCTTTCAGAGCTCCAGACCACTTACAGGCGCTGAGCTCGGCT 558
Qy      138 -ValGlnAenLeuGlnGlnSerTyMetProProAspSerSerSerProGluAsnlySse 157
Db      559 GGTTCAGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
Qy      157 r-----leuGluValSerAspThrArgPheHisSerPh 168
Db      619 GCCAGGCCAGAGAGCTCAGTGTCCCTCTGAGGAGGAGCCGCCCTCTCGTTTGTG 678
Qy      168 eSerPheTyrgIuLeuLeuYAsnValThrAsnAsnPheAspGluArgProIleSerValG 188
Db      679 GCCCTCTGTGATTTCCCGGGGACCCACAACTTCTGAGAGAG----- 724
Qy      188 yGlyAsnlySmetGlyGlnGlyGlyPheGlyValValTyTrpSglTyTrValAsnAsnTh 208
Db      725 -----CTCAAGATTCGGGAGGGGTGGCTTTGGGTGGGTGCGGTACCGGCGGTATGAGAAC 780
Qy      208 rThrValAlaValYslySleuAlaIleValMetValAspIleThrThrGlnGluLeuSgl 228
Db      781 GGTGTATCTGTGAAGAGCTGAGAGGAACCTGACTGAGTGAAGCA 840
Qy      228 nGlnPheAspGlnGluIleYsValMetAlaYsCySglnHISgluAsnLeuValGluLe 248
Db      841 GAGCTTCTGACCGAGGTGAGAGCTGTCCAGGTTTCTTCAACCAATTTGGACTT 900
Qy      248 uleuGlyPheSerSerAspGlyAspAspLeuCySleuValTyValTyMetProAsnG 268

```

```

Db      901 TGCTGGCTACTGTGCTCAGAAAGCGCTTCTACTGCTGGTGTAAGCGCTTCTCCCAACGG
Qy      268 ySerleuLeuAspArgLeuSerCys---LeuAspGlyThrProProLeuSerThrIshMe
Db      961 CTCCCTGGAGAGCCCTCTCCACTGCGCAGAACCCAGGCTGCCACCTCTCTCTGCGCTCA
Qy      287 CArgCysArgIleAlaGlnGlyAlaAlaAsnGlyIleAsnPhelLeuHISgIuAsnHIS
Db      1021 GCGACTGAGCATCTCTTGGGAGACGCCCGGCAATTCAAGTTTCTACATCAGAGACGCC
Qy      307 s-----IleHisArgAspIleLeuSerAlaAsnIleLeuLeuAspGluAlaPheThrAl
Db      1081 CAGCTCATTCATGAGACATCAAGATTCACAACTCTCTGATGAGAGCTACACC
Qy      325 aLysIleSerAspPheGlyLeuAlaArgAlaSerGlyIlePheAla-----
Db      1141 CAGCTGGAGAGCTTGGCTGCTGGCCGGTTGAGC---CGCTTGGCCGGTCCAGCCCGCAG
Qy      341 -GlnThrValMetThrSerArg-----IleValGlyThrThrAlaIleYrMetAlaPr
Db      1198 CCAAGACGACATGTGGCCCGGACACAGAGTGGCGGACCTGGGCTACCTGCCGA
Qy      357 oGluAlaLeuArg---GlyGluIleThrProLysSerAspIleYrSerPheGlyValIa
Db      1258 GGAATACATCAAGAGGAGAGGCTGCTGTGACACGACACCTTACCTTGGGCTGT
Qy      376 lLeuLeuGluIleIleThrGlyLeuProAlaValAspGluHis---ArgGluProGlnLe
Db      1318 AGTGTACAGACCTTGGCTGCTGACAGGGCTGTGAAAGACGACGCTGCAGAGACCAAGTA
Qy      395 uLeuLeuAspIleLysGluGluIleGluAspGlu-----G
Db      1378 TCTGAAGAAGCTCTGTGAGAGAGAGGCTGAGAGGCTGAGTGGCTTGAGAGACCA
Qy      407 uLysThrIleGlu-----
Db      1438 GAGACACATGCAAGAGAGTGTGGCTGACAGATGCTGGGCTGCTCCATCGGCATGACAT
Qy      412 -----AspTyrIleAspLysMetAsnAspAlaAspSerThrSerValGluAlaMe
Db      1498 CTACAGAAGACACCTGGAGCCCGGCGCTGCCACCTGAGCTGGGCT
Qy      429 tTyrSerValAlaSerGlnCysLeuHisGluLysAsnLysArgProAspIleLysLy
Db      1558 GGGCCAGCTGGCTGCTGCTGCTGACACCGCGGCCAAAGAGGCTCTCTATGACCA
Qy      449 sValGlnGlnLeuLeuGlnGlnMetThrAla
Db      1618 GGTGTACGAGAGCTAGAGAAAGCTGCAGGCA

```

```

; LENGTH: 2361
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1673

Alignment Scores:
Pred. No.: 7,96e-47 Length: 2361
Score: 508.50 Matches: 124
Percent Similarity: 55.79% Conservative: 64
Best Local Similarity: 36.80% Mismatches: 118
Query Match: 21.31% Indels: 31
DB: 10 Gaps: 9

US-10-001-254-16 (1-460) x US-09-938-842A-1673 (1-2361)

Qy      141 LeuGluGlnSerTyrMetProProAspSerSerProGluAsnLysSerLeuGluVal
Db      1294 CTCGAAAGAAAGAAAGAGCTTTCAGATGACACAGCTCCACCAAGTTGCCGTAGAGAT
Qy      161 SerAspThrArgPheHisSer-----PheSerPhe
Db      1354 GTTGAACAAGCTAAACATTCAGATCATTCCTCGAAGAAAGATAGGTTGCTTAT
Qy      171 TyrGluLeuLysAsnValThrAsnAsnPheAspGluArgProLysSerValGlyAsn
Db      1414 TTCAGAGTTCAGAGATGACCAATTAATCTTCAG-----AGA
Qy      191 LysMetGlyGluGlyGlyPheGlyValIleYrLysGlyTyrValAsnAsnThr---Thr
Db      1450 GTTCTGTGTGAAGAGAGCTTGGAGTGTGATTCATGATGATGTTATGTTATGTTAGTACCA
Qy      210 ValAlaValLysLysLeuAlaAlaMetValAspIleThrThrGluGluLeuGlnGln
Db      1510 GTACCTGTAAATTGCTCTCTCA-----TCATCTTCCCAAGCTATAAACAT
Qy      230 PheAspGlnGluIleLysValMetAlaLysCysGlnHisGluAsnLeuValGluLeu
Db      1558 TTCAAGACAGAGAGTGAACCTTCTTATGAGAGTACACATATAAATTGGTAGCTTGT
Qy      250 GlyPheSerSerAspGlyAspAspLeuCysLeuValIleYrValIleYrMetProAsnGlySer
Db      1618 GGTATTGTGATGAAGAGACCATTTGGCCCTCATACAGTACAGTCAATGAGAGAC
Qy      270 LeuLeuAspArgLeuSerCysLeuAspGlyThrProProLeuSerThrPheMetArgCys
Db      1678 TTAACAACAACATTTATAGAGAAAGCGTGGATTTGCTTAAAGCTGGAAAGTAGACTA
Qy      290 LysIleAlaGlnGlyAlaAlaAsnGlyIleAsnPhelLeuHis-----GluAsnHis
Db      1738 AGAGTACTGTCGATGACAGACATRAGTTTGAGTACTTACACACTGATGCAACCA
Qy      307 HisIleHisArgAspIleLysSerAlaAsnIleLeuLeuAspGluAlaPheThrAlaLys
Db      1798 ATGTTTCACAGAGATATTAAGAACATTAATCTTGGATGAGACCTTCCAAAGCCAA
Qy      327 lLysSerAspPheGlyLeuAlaArgAlaSerGlyLysPheAlaGlnThrValMetThrSer
Db      1858 TTACCCGATTTTGGCTTGGAGATCTTTCTCAACCAAGAAACGAAACATGTTTCACT
Qy      347 ArgIleValGlyThrThrAlaIleYrMetAlaProGluAlaLeuArgGlyGlu---IleThr
Db      1918 GTTGTGCTGGAACCTCTGTTATCTTATCCGAGATTAATCAACAAATGGTTGACA
Qy      366 ProLysSerAspIleYrSerPheGlyValValLeuLeuGluIleIleThrGlyLeuPro
Db      1978 GAGAAAGATGATGTTTACATTCGGAATGTTGATGAGATCATCAACAAACCGGCT
Qy      386 AlaValAspGluHisArgGlu---ProGlnLeuLeuLeuAspIleLysGluGluIleGlu
Db      2038 ATATATTAGCAATCTCGTGAAGAGCTCAGCTGTAGATGATGGTGGATTTATGTAGAGA
Qy      405 AspGluGluLysThrIleGluAspTyrIleAspLysLysMetAsnAspAla---AspSer

```

```
Db      2098  ACGGAGAT-----ATTGGAATATTGTTGATCCAAACCTTCACGGAGCTTACGACGTC 2151
Qy      424  ThrSerValGluAlaMetTyrSerValAlaSerGlnCysLeuHisGluIlySerAsnLys 443
          |||||
          ::::|
Db      2152  GGTTCGTCTCGAAGGCTATTGAAGCTAGCTATGTCATGTGTGAATATTTCTTCGGCAAGA 2211
          |||||
          ::::|
Qy      444  ArgProAspIleLysValGlnGlnLeuLeuGlnGluMetThrAlaSer 460
          |||||
          ::::|
Db      2212  AGACCAAGCATGTCTCAAGTTGTTAGGATCTTAAGAGTGTGTGATATCT 2262
```

Search completed: January 18, 2004, 09:06:01
Job time : 721.136 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 18, 2004, 01:11:43 / Search time 4491.9 seconds
(without alignments)
2488.940 Million cell updates/sec

Title: US-10-001-254-16
Perfect score: 2386
Sequence: 1 MNKRIPSTYVRCINVLIR.....KNKRPIKVKYQLIQEWMTAS 460

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n model -DEV=x1h
-Q/cgrr2_1/USPFO.spool/US10001554/runat_16012004_152424_19723.apd_query.fasta_1.1109
-DB=EST -QPM1=faa6ap -SUFFIX=ret -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=D10sum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10001254 @CGN 1.1 4382 @runat_16012004_152424_19723 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESOURRY -NBS -SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2017	84.5	2481	11 AK028837	AK028837 Mus muscu
2	2004	84.0	2810	11 AK029028	AK029028 Mus muscu
3	1094.5	45.9	859	10 BG164491	BG164491 602342026
4	1004.5	42.1	682	10 BF694134	BF694134 602082746
5	946	39.6	718	13 BQ780493	BQ780493 UT-R-FFO-
6	946	39.6	1161	11 AK020397	AK020397 Mus muscu
7	944	39.6	702	13 BQ782157	BQ782157 UT-R-FFO-
8	935.5	39.2	811	10 BG616438	BG616438 602642772
9	874	36.6	769	13 BU441365	BU441365 603208981
10	856	35.9	719	10 BF696981	BF696981 602130160
11	804	33.7	962	13 BU225249	BU225249 603947374
12	774.5	32.5	541	10 BG691069	BG691069 340084 BA
13	755	31.6	676	10 BB613167	BB613167 BB613167
14	752	31.5	478	12 BM431425	BM431425 IDU016D05
15	749.5	31.4	858	10 BF687921	BF687921 602066996
16	675	28.3	635	13 BQ205658	BQ205658 UT-R-EP0-
17	665	27.9	424	10 BF758197	BF758197 RC4-CT052
18	662.5	27.8	624	13 BU479576	BU479576 602472624
19	640.5	26.8	508	10 BE479760	BE479760 164577 BA
20	639	26.8	663	14 BY726858	BY726858 BY726858
21	632	26.5	557	14 CB475590	CB475590 jns109 FO
22	631	26.4	637	10 BB613447	BB613447 BB613447
23	629	26.4	576	10 BF238344	BF238344 601904613
24	627	26.3	638	14 BY721552	BY721552 BY721552
25	623	26.1	613	13 BU316755	BU316755 603850730
26	612	25.6	598	13 BQ552228	BQ552228 H4014C09-
27	597	25.0	1008	13 BX283594	BX283594 BX283594
28	588	24.6	508	12 BM918155	BM918155 AGENECOURT
29	587	24.6	610	10 BB660378	BB660378 BB660378
30	580	24.3	499	13 BU759126	BU759126 UT-R-FFO-
31	577.5	24.2	852	13 BU209111	BU209111 603950834
32	568.5	23.8	710	14 CA054774	CA054774 sea1rpb51
33	563	23.6	335	13 BU430602	BU430602 UT-HF-BNO
34	557	23.3	402	10 BE482619	BE482619 168463 BA
35	556	23.3	823	13 BU246150	BU246150 603782255
36	551.5	23.1	555	13 BQ617704	BQ617704 faa67c09.
37	549	23.0	575	4 BX522921	BX522921 RZPD Mus
38	528	22.1	503	14 CA538859	CA538859 CO272B01-
39	523	21.9	851	13 BU255115	BU255115 603746394
40	522	21.9	773	13 BU240307	BU240307 603321515
41	511.5	21.4	629	9 AL647125	AL647125 AL647125
42	503.5	21.1	804	14 CB659049	CB659049 OSJNEC15K
43	488.5	20.5	505	10 BE482323	BE482323 168064 BA
44	487.5	20.4	2675	11 AK045958	AK045958 Mus muscu
45	486.5	20.4	1434	9 AI967314	AI967314 Ljirpesc

ALIGNMENTS

RESULT 1
AK028837 2481 bp mRNA linear HTC 05-DEC-2002
LOCUS
DEFINITION Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:473246109 product:interleukin-1 receptor-associated kinase 4 [Mus musculus], full insert sequence.
ACCESSION AK028837
VERSION AK028837.1 GI:26324783
KEYWORDS HTC, CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

- REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCES
- 1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
- 2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
2049374
11042159
- 3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system -384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
- 4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kuchiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schiraldi, L.M., Staudt, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaur, P., Nordone, P.,
Rising, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
- 5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2481)
- TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCES
AUTHORS
- ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., BONO, H., CARNINCI, P.,
FUKUDA, S., FURUNO, M., HANAGAKI, T., HARA, A., HASHIZUME, W.,
HAYASHIDA, K., HAYATSU, N., HIRAMOTO, K., HIRAKA, T., HIZOANE, T.,
HORI, F., IMOTANI, K., ISHII, Y., ITOH, M., KAGAWA, I., KASUKAWA, T.,
KATO, H., KAWAI, J., KOJIMA, Y., KONDO, S., KONO, H., KODA, M.,
KOYA, S., KURIHARA, C., MATSUYAMA, T., MIYAZAKI, A., MURATA, M.,
NAKAMURA, M., NISHI, K., NOMURA, K., NUMAZAKI, R., OHNO, M., OHSAKI, N.,
OKAZAKI, Y., SAITO, R., SAITOH, H., SAKAI, C., SAKAI, K., SAKAZUME, N.,
SANO, H., SAKAI, D., SHIBATA, K., SHINAGAWA, A., SHIRAKI, T.,
SOGBE, Y., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TAKAKU-AKAHIRA, S.,
TAKEDA, Y., TANAKA, T., TOMATSU, A., TOYA, T., YASUNISHI, A.,
MURAMATSU, M. and HAYASHIZAKI, Y.

COMMENT
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216]
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES
source
1..2481
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:4732460109"
/db_xref="taxon:10090"
/clone="4732460109"
/tissue_type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
116..1495
/note="unnamed protein product; interleukin-1
receptor-associated kinase 4 [Mus musculus]
putative."

CDS
650 a 592 c 634 g 605 t
ORIGIN
BASE COUNT
1 MetAsnLysPrcIleThrProSerThyValArgCysLeuAsnValGlyLeuileArg 20
116 ATGACCAACCGCTGTGACACCATCGACATCATACCACTTAATGTGGGATCTTAGG 175
21 LysLeuSerAspPheLeuAspProGlnGluGlyTTPpLysLeuAlaValAlaLeuys 40
176 AAGCGTTCGATTTATATGATCTTCAGAGAGGCTGGAAGAAATVAGCATCTACTACAA 235
41 LysPrcSerGlyAspAspArgTyrAsnGlnPheHisIleArgArgPheGlnAlaLeu 60
236 AAGCCGTCGGGACGACACGATACATCATCGTTCCATATAGAGAGATTCAGACCTTACTT 295
61 GlnThrGlyLysSerProThrIserGlnLeuLeuPheAspTyrGlyThrThraAsyThr 80
296 CAGACCGGGAAGAGCCCACTGTGACTGCTTTGACTGGGGACCAAGCAACTGCACA 355
81 ValGlyAspLeuValAspLeuLeuileGlnAsnGlnPheAlaProAlaSerLeuLeu 100
356 GTTGGGACCTTGATCTACTGTCAGATGAGCTGTGGCTTGGCCCGGACCTCTCGG 415

US-10-001-254-16 (1-460) x AK028837 (1-2481)

Alignment Scores:
Pred. No.: 1,18e-221 Length: 2481
Score: 2017.00 Matches: 385
Percent Similarity: 90.85% Conservative: 32
Best Local Similarity: 83.88% Mismatches: 42
Query Match: 84.53% Indels: 0
DB: 11 Gaps: 0

QY 101 LeuProAaPalaValProLeuThrAlaAsnThrLeuProSerLeuGluAlaIleThrVal 120
 Db 416 CTGGCCGATGCCGTTCCCAAAACCGTCAAAAGCGCTCCCTAGAGAAAGCGGCAACAGTG 475
 QY 121 GlnGlnLeuSerProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140
 Db 476 GCACAAACACACCGGCGCTTGTCCAGAAAAAGACAGACATCCGTAATCCCTATGCCGCAAG 535
 QY 141 LeuGlnLeuSerThrMetProProAspSerSerProGluAsnLysSerLeuGlnVal 160
 Db 536 CTAAGAACACAGCTCGAGCCACCGAGCTCTCAAGCCACGAAAGAGTGAAGTCC 595
 QY 161 SerAspThrArgPheHisSerPheSerPheThrGluLeuLysAsnValThrAsnAspPhe 180
 Db 596 AGCCACACTCGGTTCCACAGCTTCCGTTCCATGAATGAAGAGATGACAAACACTTC 655
 QY 181 AspGluArgProIleSerValGlyGlyAsnLysMetGlyGlnGlyGlyPheGlyValVal 200
 Db 656 GACAGACAAACCGCGCTCGCGGTGGCAACCGATGGAGAGGGGGGATTTGGAGTGGTG 715
 QY 201 TyrIleGlyThrValAsnAsnThrThrValAlaValLysLysLeuAlaIleMetValAsp 220
 Db 716 TACAAAGGCTGTGTAACACACCACTGTGGCGGTGAAGAAAGCTCGAGCGATGGTTGAA 775
 QY 221 IleThrThrGlnGluLeuLysGlnGlnPheAspGlnGlnIleLysValMetAlaLysCys 240
 Db 776 ATCAGTACTGAAGAACTTAAGCAACAGTTGATCAAGAAATTAAAGTAAATGGCAACGTGT 835
 QY 241 GlnHisGluAsnLeuValGlnLeuLeuGlyPheSerSerAspArgLysAspLeuCysLeu 260
 Db 836 CAGACACAGAACCTCGTGAGCTCTCGGCTTCTCCAGCGACACGCAACCTGTGCTTA 895
 QY 261 ValTyrValTyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThr 280
 Db 896 GTGATGCTTACATAGCCCAACGGGTCCTGTGAGACAGACTGCTCTCGTGAATGATGACA 955
 QY 281 ProProLeuSerThrPheHisMetArgCysLysLysIleAlaGlnGlyAlaAlaAsnGlyIleAsn 300
 Db 956 CCACCGCTTCTCGGCACACAAAGGTGGACAGGTGCTCTCAGGGAGACGCAAAATGGATCAGG 1015
 QY 301 PheLeuHisGluAsnHisIleHisArgAspIleLysSerAlaAsnIleLeuLeuAsp 320
 Db 1016 TTTCTGATGATGAATCATCATTCATTCATGACATTTAAAGTCAAAATATCTTACTGAC 1075
 QY 321 GlnAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGluLysPheAla 340
 Db 1076 AAGACATTACTGCCAAAATATCTGACTTGTGGCTTGCACGGGCTTCCGCAAGGCTAGCG 1135
 QY 341 GlnThrValMetThrSerArgIleValGlyThrThrAlaTyrMetAlaProGluAlaLeu 360
 Db 1136 CAGACGGTCAATGACACCGCAATGCTGGGCAACCGCTTACATGGCACCCGAAGCTTTG 1195
 QY 361 ArgGlyGluIleThrProLysSerAspIleTyrSerPheGlyValIleLeuLeuGluIle 380
 Db 1196 CGGGGAAATTAACACCAATCTGACATCTACAGCTTCCGGCTGCTTCTGTGAGACTG 1255
 QY 381 IleThrGlyLeuProAlaValAspGluHisArgGluProGlnLeuLeuAspIleLys 400
 Db 1256 ATAACCGGCTGGCGGTGTGATGAAGAAACCGTAAACCTCAACTACCTGGATATTA 1315
 QY 401 GlnGluIleGluAspGlnGluLysThrIleGluAspTyrIleAspLysMetAsnAsp 420
 Db 1316 GAAGAGATTGAAGTGAAGAGAACATTCATTAACGATGAGATGAGAGATAGCGAT 1375
 QY 421 AlaAspSerThrSerValGluAlaMetTyrSerValAlaSerGlnCysLeuHisGluLys 440
 Db 1376 GCGGACCTCGCTCGGTGGAAGCATATCTCTCTGACCGCATGTCTGATGAGAGAG 1435
 QY 441 LysAsnLysArgProAspIleLysLysValGlnGlnLeuGlnGlnIleMetThrAla 459
 Db 1436 AAAACACAGACGGCCAGCATTTGCAAGGTTCAACAGCTGCTACAAAGAGATGTCTGCT 1492

AK029028
 LOCUS
 DEFINITION
 Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:4732482P03 product:interleukin-1 receptor-associated kinase 4 [Mus musculus], full insert sequence.
 AK029028
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P., and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishize, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubara, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, D., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichipillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamazaki, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staib, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Guestinch, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohetsuki, S. and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11217851
 5
 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6
 (bases 1 to 2810)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hirozane, T.,

Horii F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kanakawa T., Katoh H., Kawabi J., Kojima Y., Kondo S., Konno H., Koude M., Koya S., Kurikawa C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi R., Nomura K., Nunazaki R., Ohno M., Ohsato N., Okazaki Y., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shitagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akihira S., Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M. and Hayashizaki Y.

TITLE
Direct Submission

JOURNAL
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1 .2810
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:4732482P03"
/db_xref="taxon:10090"
/clone="4732482P03"
/tissue_type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
176 .1556
/note="interleukin-1 receptor-associated kinase 4 [Mus musculus]
putative"

misc_feature

BASE COUNT	738 a	677 c	715 g	680 t
ORIGIN				

Alignment Scores:

Pred. No. :	4,47e-220	Length:	2810
Score:	2004.00	Matches:	385
Percent Similarity:	90.65%	Conservative:	32
Best Local Similarity:	83.70%	Mismatches:	42
Query Match:	83.99%	Indels:	1
DB:	11	Gaps:	0

```

US-10-001-254-16 (1-460) x AK029028 (1-2810)
QY      1 MetaAmlyPrIoleThrProSerThrTYValArgCYLeuAsnValGIyNeuileArg 20
       176 ATGAACAAGCGCTTGACACCATCGACATACATACCAACGAACTTAATGTGGGATCCTTAGG 235
QY      21 LysLeuSerASPheIlleaspProGIngluglYTrrPLyLysLeuAlValAlalleys 40
       236 AAGCGTCCGATTATTCATCCTCAAGAAGGCTGAGAAATAATTGCACTACTACTCAAA 295
QY      41 LysProSerGLYAspAspArgTYrAngInPheHisIleArgArXPheGlulaleuLen 60
       236 AAGCGTCCGCGGACACAGATACATTCAGTTCCATATAGAGATTGCAACCCTTACTT 355
Db      356 CAGACCGGGAAGGCCCACTGTGAACGCTGTTGACTGGGGGACACCAAGATGCACA 415
QY      61 GlnrhngLYLyeSePrOthIsErGIueLUeUpeAsPTrrgLYThrThaAnCYsthr 80
       356 CAGACCGGGAAGGCCCACTGTGAACGCTGTTGACTGGGGGACACCAAGATGCACA 415
Db      81 ValGIyAspLeuValAspleuleuileGInanGluPhePhalAProAlSerleuLen 100
       416 GTTGGCGAAGCTTGATCTACTGTGGCGAATTTGAGCTGTTGCCCCCGCACTCCG 475

```

QY	101	LeuProAspAlaValProlySerThrAlaSerThrLeuPProSerIleuValAlaIleThrVal	120
Db	476	CTCCGGGATGCCGTTCCCAAAACCGTCAAAAAGCTGCTCCTTAAGAAAGCGGCAACGTG	535
QY	121	GlAlGlnLysGlnMetProPheCysAspIlyAspArgThrLeuMetThrProValGlnAsn	140
Db	536	GCACAAACACACGGGCGCTTGTCAAGAAAAAGGACAGGACATCCGTAAATGCTCATGCCAG	595
QY	141	LeuGlnGlnSerTyThrMetProProAspSerSerSerProGluAsnLysSerLeuVal	160
Db	596	CTGAAGACACAGCGTCGACGACCCAGGACCTCTCAAGCCAGACCAACAGAAAGTGTAGAGTCC	655
QY	161	SerAspThrArgPheHisSerPheSerPheTyArgIleuLysAsnValThrAspAsnPhe	180
Db	656	AGGACACCTCGGTTCCACAGCTTCTCGTCCATGAACTGAAAGAGCATCAACAAACATTC	715
QY	181	AspGlnArgProIleSerValGlyIleAsnLysMetGlyGlnGlyPheGlnValVal	200
Db	716	GACGACGACACCGGCGTCTCCGGTGGCAACCGAGTGGGAGAGGGGGATTTGGAGTGGTG	775
QY	201	TyTlySerGlyTyThrValAsnAsnThrThrValAlaValLysLysLeuAlaAlaMetValAsp	220
Db	776	TACAAAGGCTGTGTGAACAACACCATCGTGGCGGTGAAGAAAGCTCCGAGCATGTTGAA	835
QY	221	IleThrThrGlnGlnIleuLysGlnGlnPheAspGlnGlnIleLysValMetAlaLysCys	240
Db	836	ATCAGTATCTGAAGAAATTAAGAACAGATTGTATCAAGAAATTAAGATTAATGGCAACGTG	895
QY	241	GlnHisGlnLysAsnLeuValGlnLeuLeuGlnGlyPheSerSerAspArgLysAspAspLeuCysLeu	260
Db	896	CAGCAGCAAGAACCTGGTGGAGCTGCTCGGCTTCTCCACGACAGGACGACAACTGTGCTTA	955
QY	261	ValTyThrValTyThrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThr	280
Db	956	GTCGATGCTTACATGCCCAACGGGTCCTTGCTGCAACAGCTGCTCGCTCGATGGATGATCA	1015
QY	281	ProProLeuSerThrPheIleMetArgCysLysIleAlaGlnIleAlaAlaAsnGlyIleAs	300
Db	1016	CCACCGGTTTCTGGGACAGCAAGAGTGCACAGGTTGCTAGGGGACACGCAAAATGGCATGAC	1075
QY	300	nPheLeuHisGlnAsnHisHisIleHisArgAspIleLysSerAlaAsnIleLeuLeuAs	320
Db	1076	GTTTCTGCATGAAATCATCATCTCATAGATTTTAAAGTGCAAATATCTTACTAGA	1135
QY	320	ArgAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGlnLysPheAl	340
Db	1136	CAAAAGCTTATCGCAAAATATCTGACTTGGGCTTGCACGGGCTTGGCAAGGCTAGC	1195
QY	340	ArgIleThrValMetThrSerArgIleValGlyThrThrAlaTyThrAlaProGluAlaLe	360
Db	1196	GCGACGCGCTATACCAAGCCGATCGTGGGCAACACGGCTTACATCGGACCCGAAAGCTTT	1255
QY	360	ValArgGlnGlnIleThrProLysSerAspIleTySerPheGlyValValIleLeuGlnIle	380
Db	1256	GCGGGGGAATATACCCCAATCTGCATCTTACAGCTTCGGCGTGGTTCCTTGGAGCT	1315
QY	380	eIleThrGlyLeuProAlaValAspGlnHisArgGlnProGlnLeuLeuAspIleLys	400
Db	1316	GATACCGGCGCTGGCGCTGTGGATGAAAAACCGTAACCTCAACTACTCGTGAATATTAA	1375
QY	400	ArgIleGlnIleGlnAspGlnGlnLysThrIleGlnLysPtyrIleAspLysLysMetAsnAs	420
Db	1376	AGAAGAGATTGAAGAGAAGAAGACGATTGAAATTAACACGGATGGAAGATGAAGCGA	1435
QY	420	ProLysAspSerThrSerValGlnAlaMetCysSerValAlaSerGlnCysLeuHisGlnLys	440
Db	1436	TGCGGACCTGCTTCGTGTGAAGCAATGACTCTCTCTGACCAAGTCTTGCATGAGAA	1495
QY	440	SlyAsnLysArgProAspIleLysLysValGlnIleuLeuGlnGlnIleMetThrAla	459
Db	1496	GAATAACAGACGGCGACATCTTGCAAGGTTCAACAGCTGCTTACAAAGAAAGTCTGCT	1553

BG164491
 LOCUS 859 bp mRNA linear EST 06-FEB-2001
 DEFINITION 602342026P1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4452055 5',
 mRNA sequence.
 ACCESSION BG164491
 VERSION BG164491.1 GI:12671194
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 859)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LHM10240 row: d column: 08
 High quality sequence stop: 634.
 Location/Qualifiers
 1..859

FEATURES

source
 1..859
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4452055"
 /tissue_type="hypertrophoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_89"
 /note="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: NCI;
 Site 2: Sall; Cloned unidirectionally; Oligo-dT primed.
 Average insert size 1.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 266 a 168 c 190 g 235 t
 ORIGIN

Alignment Scores:

Pred. No.: 1 23e-115 Length: 859
 Score: 1094.50 Matches: 228
 Percent Similarity: 91.05% Conserved: 6
 Best Local Similarity: 88.72% Mismatches: 19
 Query Match: 45.87% Indels: 8
 DB: 10 Gaps: 2

US-10-001-254-16 (1-460) x BG164491 (1-859)

QY 1 MetAsnlysProIleThrProSerThrTyValArgCysLeuasnValGlyLeuIlearg 20
 |||||
 DB 35 ATGACCAACCAATCAACCATCAATATGTGGCTGCTCAATGTGACATTAATAGG 94
 |||||
 QY 21 LysLeuSerAspPheIleAspProGlnGlyGlyTrpLysLysLeuAlaValAlaIleLys 40
 |||||
 DB 95 AACCTGCACATTTTATGATCTCCAGAGAGATGAGAAAGATTAGCTGTAGCTATTTAA 154
 |||||
 QY 41 LysProSerGlyAspAspArgTyraGlnPheHisIleArgArgPheGlnAlaLeuLeu 60
 |||||
 DB 155 AAACCATCTGTGATGATGATACATCACTGTTCACTTAAGAGATTGAAGCACTTACTT 214
 |||||
 QY 61 GlnThrGlyLysSerProThrSerGlnLeuLeuPheAspTrpGlyThrThrAsnCysThr 80
 |||||
 DB 215 CAACACTGGAAGAAAGTCCCATCTTCTGATTAATCTGTTGACCTGGGACACCAAAATTGACA 274
 |||||
 QY 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPhePheAlaProAlaSerLeuLeu 100
 |||||
 DB 275 GTTGGTATCTTGTGATCTTTTGTATCCAAAGAAATTTTGTCTCTGCAAGTCTTTTG 334

QY 101 LeuProAspAlaValPProLysThrAlaAsnThrLeuProSerLysGlnAlaIleThrVal 120
 |||||
 DB 335 CTCCCAATGTCTGTCTCCCAAAAGCTTAATACATCTTCTTAAGAGCTAATACAGTT 394
 |||||
 QY 121 GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140
 |||||
 DB 395 CAGCAAAAACAGATGCTCTTCTGTGACAAAGACAGACATTTAGACACCTGTGACAGAT 454
 |||||
 QY 141 LeuGlnGlnSerTyMetProProAspSerSerSerProGlnLeuLysSerLeuGlnVal 160
 |||||
 DB 455 CTTCGACAAAGCTAATATGACCACTGCTCTCAAGTCCAGAAAATATAAGTTTGAAGATT 514
 |||||
 QY 161 SerAspThrArgPheHisSerPheSerPheThrGlnLeuLysAsnValThrAsnAspHe 180
 |||||
 DB 515 AGTATACACGCTTTTCAAGTTTTCATTTTGAATTGAAGATGTCACAAATAACTTT 574
 |||||
 QY 181 AspGlnArgProIleSerValGlyLysAsnLysMetGlyLysGlyPheGlyValVal 200
 |||||
 DB 575 GATGAACGACCACTTTCTGTTGGTGTAATTAATGAGAGAGGAGATT-GGAGT-GTA 632
 |||||
 QY 201 TyrlsGlyTyTyValAsnAsnThrThrValAlaValLysLysLeuAlaIleMetValAsp 220
 |||||
 DB 633 TATTAAGGCTACGTAATTAATACACAGAGTGTGAGTAAGAACTTGACAG-ATGTTGAC 691
 |||||
 QY 221 IleThrThrGlnGlnLeuLysGlnGlnPheAspGlnGlnLysValMetAlaLysCys 240
 |||||
 DB 692 CTTCACCTGAGAGAACTGAACAGCGTTTTCACAGA--CTTCACTTAATGGCCAG-TGT 747
 |||||
 QY 241 GlnHisGlnAsnLeuValGlnLeuLeuGlyPheSerSerAspGlyAspAsp 257
 |||||
 DB 748 TCACCTGAAGAACTGTGAACCC-----TGGTCTTCAGAGAGCGGATTAAC 789

RESULT 4
 BF694134
 LOCUS 682 bp mRNA linear EST 22-DEC-2000
 DEFINITION 602082746P1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4247171 5',
 mRNA sequence.
 ACCESSION BF694134
 VERSION BF694134.1 GI:11979542
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 682)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: CLONETECH Laboratories, Inc.
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LHM1063 row: k column: 12
 High quality sequence stop: 617.
 Location/Qualifiers
 1..682

FEATURES

source
 1..682
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4247171"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_id="NIH_MGC_81"
 /note="Organ: muscle (skeletal); Vector: pDNR-LIB
 (cloneech); Site 1: SfiI (ggccatcggcc); Site 2: SfiI
 (ggccatcggcc); 5' and 3' adaptors were used in cloning
 as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCGC-3'
 and 3' adaptor sequence:
 5'-ATTCTAGAGCGGAGCGGCGGACGATG-dT(30)BN-3' (where B = A,
 C, G, or T).

C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA).

BASE COUNT 222 a 119 c 142 g 198 t 1 others

ALIGNMENT SCORES:

Pred. No.:	Length:	Matches:
Score: 2.26e-105	682	212
Percent Similarity: 82.81%	Conservative: 0	
Best Local Similarity: 82.81%	Mismatches: 12	
Query Match: 42.10%	Indels: 34	
DB: 10	Gaps: 1	

US-10-001-254-16 (1-460) x BF694134 (1-682)

```

QY 122 GlnUleuGlnMeProPhCyAspArgThrLeuMetThrProValGlnAanLeu 141
DB 1 CAAAAACAGATGCTTCTTCTGTGCAAGACAGACATTATGACACTGTGCAGAACTTT 60
QY 142 GlnUlnSerTyMeCProAspSerSerProGlnAanUlySerLeuGlnValSer 161
DB 61 GAACAAAGCTATATGCACTGACCTCTCAAGTCCAGAAAATAAAGTTTAAAGAGTACT 120
QY 162 AspThrArgPheHisSerPheSerPheTyGlnUleuLysAsnValThrAsnProPheAsp 181
DB 121 GATACACGTTTACAGTGTTCATTTTAAAGATGAAATGTCACAAATTAACCTTTGAT 180
QY 182 GlnArgProLysSerValGlyGlyAanUlyMetGlyGlnUlyGlyPheGlyValValTy 201
DB 181 GAACGACCACTTTCT----- 195
QY 202 LysGlyTyValAanAsnThrThrValAlaValLysLysLeuAla-AlaMetValAsp11 221
DB 196 -----GNTGCTGATGCTTGACAT 213
QY 221 eThrThrGlnUlnUleuLysGlnInPheAspGlnUlnUlyValMetValLysCysG1 241
DB 214 TACTACTGAAAGCACTGAAACAGAGTTTATGATCAAGAAATTAAGTAAATGGCAAGTGTCA 273
QY 241 nHisGlnAanUleuValGlnUleuLysGlyPheSerSerAspGlyAspAspLeuUlyVal 261
DB 274 ACATGAAACTTATGTAAGTACTTGTCTTCTAAGTATGAGTATGACCTGCTTACTAGT 333
QY 261 lTyValTyMeCProAsnGlySerLeuLeuAanPargLeuSerCysLeuAspGlyThrPr 281
DB 334 ATATGTTTACATGCTTAATGTTTCATTTGCTAGACAGACTCTCTTGTGATGATGCTTCTCC 393
QY 281 oProLeuSerThrHisMetArgCysLysLysLeaGlnGlyAlaAlaAsnGlyLysAsnPh 301
DB 394 ACCACTTCTTGGCAGATGAGATGCAAGATTTGCTACAGGTCAGCTAATGTCATCAATTT 453
QY 301 eLeuHisGlnAanHisHisLeuHisArgAspLysSerAlaAsnUlyLeuLeuAspG1 321
DB 454 TCTACATGAAATCAATCATATTCATAGACATTAATAAGTCAGAAATTTCTTACTGATGA 513
QY 321 uAlaPheThrAlaLysLysSerAspPhe-GlyLeu-AlaArgAlaSerGlyUlyPheAla 340
DB 514 AGCTTTTACTGCTTAATAATATCTGACTTGTGCGCTTTCACAGGCTTCTGAGAACTTGGC 573
QY 341 GlnThrValMetThrSerArgLysGlyValThrAlaTyMetAlaProGlnUlnAlaLeu 360
DB 574 CAAACAGTCAATGCTGACAGAAATGTCGGAACAACACT-TATATGGAGCAGCAAGCTTTG 632
QY 361 ArgGlyGlnUlyLeuThrProLysSerAspLysTySerPheGly 374
DB 633 CGTGGAGAAATTAACCAAA-TCTGATATTAACAGCTTGGGG 673

```

RESULT 5
BQ780493/c 718 bp mRNA linear EST 26-JUL-2002

DEFINITION UI-R-PF0-coz-d-10-0-UI-s1 UI-R-PF0 Rattus norvegicus cDNA clone
UI-R-PF0-coz-d-10-0-UI 3', mRNA sequence.
ACCESSION BQ780493
VERSION BQ780493.1 GI:21988965
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 718)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL 9704447
MEDLINE 8889548
PUBMED 8889548

COMMENT

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Jeff Stevens
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: DISTRIBUTION: Researchers may obtain clones
from Research Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES

source

location/Qualifiers
1..718
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="UI-R-PF0-coz-d-10-0-UI"
/tissue_type="Mixed tissues"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-R-PF0"
/note="Vector: pT73-Pac (Pharmacia) with a modified
polylinker; Site 1: Ecor I; Site 2: Not I; UI-R-PF0 is a
subtracted cDNA library containing the following tissues(s)
): Normal cartilage and SR-JMS Tumor Line. The
subtraction was made according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for these libraries are: CTATGAGAC,
CATTTTGA.
TAG LIB=UI-R-PF0
TAG TISSUE=cartilage
TAG_SEQ=CTATGAGAC"

BASE COUNT 158 a 174 c 160 g 225 t 1 others

ORIGIN

ALIGNMENT SCORES:
Pred. No.: 1.41e-98 Length: 718
Score: 946.00 Matches: 184
Percent Similarity: 93.24% Conservative: 9
Best Local Similarity: 88.89% Mismatches: 14
Query Match: 39.65% Indels: 0
DB: 13 Gaps: 0

US-10-001-254-16 (1-460) x BQ780493 (1-718)

```

QY 253 SerAspGlyAspAspLeuUlyValTyValTyMeCProAsnGlySerLeuAanP 272
DB 716 AGCGACAGTACAACTGTGCTTATGCTTACATGCTCCAAATAGTCTTCTTCTTACAG 657

```

Qy	273	ArgLeuSerCysLeuAspGlyThrProProLeuSerTPH1SMeCArgCysLeu1Lea1a	292
Db	656	AGACTGTCTCTCGTCGACGGAACCGCACCTTTCTTGCACTAGATGCAAGATTGCT	597
Qy	293	GlnGlyAlaAlaGlnGlyIleAsnPheLeuHISGLUaSnHISGLIleHISArgAspIle	312
Db	596	CGGGGTGAGAGAAACCGGATCCGTTTGTGATGAAATCATCATATTACAGAGATATT	537
Qy	313	LysSerAlaAsnIleLeuLeuAspGluAlaPheThrAlaValSileSerAspPheGlyLeu	332
Db	536	AAAGTCGAATATCTTCTGACCAACACTTACTGCCAATAATCTGACTTTGGGCTT	477
Qy	333	AlaArgAlaSerGlyLysPheAlaGlnThrValMetThrSerArgIleValGlyThr	352
Db	476	GCACGGGCTTCGCGAAGCTTGACAAACGGTCATGACCGCAATTTGTGGGCAACG	417
Qy	353	AlaTyrMetAlaProGluAlaLeuArgGlyGluIleThrProLysSerAspIleTyrSer	372
Db	416	GCTTATATGCACTCGAAGCTTTGCGAGAGAAATTAACACCAATCTGACATCTACAGC	357
Qy	373	PheGlyValValLeuLeuGluIleIleThrGlyLeuProAlaValaAspGluHISArgGlu	392
Db	356	TTCCGTGGTCTTTATTTAGACTATTAACCGGACTTGGCGCTGTGATGAAACCGGAA	297
Qy	393	ProGlnLeuLeuLeuAspIleLysGlnGluIleGluAspGluGlyThrIleGluAsp	412
Db	296	CCTCAACTACTACTGATATTAAGAAGAGATTGAAGCGAGAGAGACATCGAAGAT	237
Qy	413	TyrIleAspLysLysMetAsnAspAlaAspSerThrSerValGluAlaMetTyrSerVal	432
Db	236	TACACAGACGAGAGATGAGGATGCGGACCTCGCTCGGTGAGGCGGATCTACTGTT	177
Qy	433	AlaSerGlnCysLeuHISGLUaLysAlaSnLysArgProAspIleLysLysValGlnGln	452
Db	176	GCTAGCCAGTCTCTGCAAGAGAAATAACAGACGCCAGCATTTGCAAGTTCAACG	117
Qy	453	LeuLeuGlnGluMetThrAla 459	
Db	116	CTGCTACAGAGACCTCTGCA 96	
RESULT 6			
AK020397			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			

with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from lambda phage I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B

FEATURES

source

Location/Qualifiers
1..1161
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM:DB:9330209D03"
/db_xref="MGI:1911393"
/db_xref="taxon:10090"
/clone="9330209D03"
/sex="male"
/issue_type="diencephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
183..776
/note="unnamed protein product; interleukin-1 receptor-associated kinase 4 [Mus musculus]
putative"
/codon_start=1
/protein_id="BAB32090.2"
/db_xref="GI:15919912"
/db_xref="MGI:1924812"
/translation="MNKPLPSTFYIRNLNVLIRKLSDFIDPQGWKLAIAIKKPSG
DPRYNOFHIRRFEALLOTGKSPICELFDMDTNCVGLVDLLVOIELFAPATLLP
DAVPOVKSILPKKAGNSGNTNPLSGKODIRNAAKRTQRAIRGLISPDNRVSES
SDTRFHSFSEHKSTNNFDEOPASAGNRMGSGP"

BASE COUNT 317 a 296 c 308 g 240 t

ORIGIN

Alignment Scores:

Pred. No.:	2,79e-98	length:	1161
Score:	946.00	Matches:	187
Percent Similarity:	87.45%	Conservative:	22
Best Local Similarity:	78.24%	Mismatches:	29
Query Match:	39.65%	Indels:	2
DB:	11	Gaps:	0

US-10-001-254-16 (1-460) x AK020397 (1-1161)

QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
DB 183 ATGAACAGCGCTTGACACCATGACATACATACGCAACTTATGGGAGATCCTTAGG 242
QY 21 LysLeuSerAspPheIleAspProGlnGlnGlyTTrpLysLysLeuAlaValAlaIleLys 40
DB 243 AACCTGCGATTTTATGATCTCTCAAGAGGGGTGGAAGAAATTAGCATGATCATCAAA 302
QY 41 LysProSerGlyAspAspArgTyAsnGlnPheHisIleArgArgPheGluAlaLeuLeu 60
DB 303 AACCGTCCGGCAGACAGATACATCATCTTCAATTAAGAGATTCGAAAGCTTTCCTT 362
QY 61 GlnThrGlyLysSerProThrSerGlnLeuLeuPheAspTrpGlyThrThrAsnCysThr 80
DB 363 CAGACCGGGAAGAGCCCACTGTGAACCTGCTGTTTACATCGGGACACAGAACTGCACA 422
QY 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGlnPhePheAlaProAlaSerLeuLeu 100
DB 423 GTTGGCAGACTTGTGATCTACATGCTCAATGATGATGCTTTGCCCGGCACTCTCTG 482
QY 101 LeuProAspAlaValAlaPolySerThrAlaAsnThrLeuProSerLysGluAlaIleThrVal 120
DB 483 CTGCGGATGCTGTTCCCAAAACCGTCAAAAGCTGCTCTTAAGAGAGCGGCAACAGT 542
QY 120 LglnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAs 140
DB 543 GGCACAAACACAGCGGCTTGTACAGAAAAGAGACAGACATCCGTATGCTTATGCCAA 602
QY 140 nLeuGlnGlnSerTyMetProAspSerSerSerProGlnLysLysSerLeuGlnVal 160
DB 603 ACTAGAACACAGCTGCAGCCACCGGACTCTCTC-AGCCACAGACAAAGAGTGTAGATC 661

QY 160 LysAspThrArgPheHisSerPheSerPheTyGlyLeuLeuAsnValThrAsnAsp 180
DB 662 CAGCAGACTCGGTTCCACAGCTTTTCCTTCATGACGAGACATACACAAACACTT 721
QY 180 eAspGluArgProIleSerValGlyLysAsnLysMetGlyGlyGlyPheGlyValVal 200
DB 722 CCACGACAGAACCCGGCTGCGCGGTGGCAACCGGATGGAGAGGGGGATTTTGATGCT 781
QY 200 LysTyGlyGlyTyValAsnAsnThrThrValAlaLysLysLeuAlaAlaMetValAs 220
DB 782 GTACAAAGGCGCTGTGAAACACACCATGTCGCGGTGAGAAACTCGAGGCAATGCTTCA 841
QY 220 pIleThrThrGlnGlnLeuLysGlnGlnPheAspGlnGlnIleLysValMetVal 238
DB 842 ATTCAGTACTGAAGAACTTAAGCAACAGTTGATCAAGAAATTAAGTATGCA 896

CDS

RESULT 7
BQ782157/c 702 bp mRNA linear EST 26-JUL-2002
LOCUS BQ782157
DEFINITION UI-R-PF0-cp1-c-20-0-UI-81 UI-R-PF0 Rattus norvegicus cDNA clone
UI-R-PF0-cp1-c-20-0-UI 3', mRNA sequence.
ACCESSION BQ782157
VERSION BQ782157.1 GI:21990629
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 702)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLES Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Jeff Stevens
CDNA library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
DNA Distribution: DISTRIBUTION: Researchers may obtain clones
from Research Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

Location/Qualifiers
1..702
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="UI-R-PF0-cp1-c-20-0-UI"
/issue_type="Mixed tissues"
/dev_stage="Adult"
/lab_host="PH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-R-PF0"
/note="Vector: pT73-Pac (Pharmacia) with a modified polylinker. Site 1: EcoR I; Site 2: Not I; UI-R-PF0 is a subtracted cDNA library containing the following tissues (1): Normal cartilage and SR-DMS Tumor Line. The Soares, Genome Research, 6:791-806, 1996. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tags for these libraries are: CTAATGACG, CATCTGTGA.
TAG_LIB=UI-R-PF0

TAG TISSUE=cartilage
 TAG_SEQ=CTAATGACG
 BASE COUNT 158 a 169 c 156 g 219 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.33e-98 Length: 702
 Score: 944.00 Matches: 183
 Percent Similarity: 94.55% Conservative: 8
 Best Local Similarity: 90.59% Mismatches: 11
 Query Match: 39.56% Indels: 0
 DB: 13 Gaps: 0

US-10-001-254-16 (1-460) x BQ782157 (1-702)

```

Qy 258 LeuCyLeuValTyrValTyrMetProAsnGlySerLeuLeuAspArgLeuSerCysLeu 277
    |||
Db 701 CTGTCCTTAGTGAAGCTTACATGCTCCCAATGGTTCCTTCTAGACAGACTGTCTGCTGCTG
    |||

Qy 278 AspGlyThrProProLeuSerTyrPheMetArgCysLysIleAlaGlnGlyAlaAlaAsn 297
    |||
Db 641 GACGGACGCCACACCTTCTCTGACATGATGCAAGATTGCTCGGGTGACGCAAC 582
    |||

Qy 298 GlyIleAsnPhelLeuHisGluAsnHisIleHisArgAspIleLysSerAlaAsnIle 317
    |||
Db 581 GGCATCCGTTTGTGATGAAATCATATTCACAGAGATATTAAAGTGCAATATC 522
    |||

Qy 318 LeuLeuAspGluAlaPheThrAlaValIleSerAspPheGlyLeuAlaArgAlaSerGlu 337
    |||
Db 521 TTACTACACAGACTTACTGCTCAAAATATCTGACTTGGCTTGCACGGCTTCGCG 462
    |||

Qy 338 LysPheAlaGlnThrValMetThrSerArgIleValGlyThrThrAlaTyrMetAlaPro 357
    |||
Db 461 AAGCTTGACAAACGGTCATGACACCGCAATTTGGGACACAGCTTATATGGCACT 402
    |||

Qy 358 GluAlaLeuArgGlyGluIleThrProLysSerAspIleTyrSerPheGlyValValLeu 377
    |||
Db 401 GAAGCTTGGCAGAGAAATACACCCAAATCTGACATCAGCTTCGGGTGGTTT 342
    |||

Qy 378 LeuGluIleIleThrGlyLeuProAlaValAspGluHisArgGluProGlnLeuLeu 397
    |||
Db 341 TTGAGAGTGAATACCGGACTTGGCTGGATGATAAAACCGCAACTCACTACTG 282
    |||

Qy 398 AspIleLysGluGluIleGluAspGluGlyLeuThrIleGluAspTyrIleAspLys 417
    |||
Db 281 GATTTTAAAGAGATGATTGAAGCAGAGAGAGAGACGATCGAAGATTTCACAGACGAG 222
    |||

Qy 418 MetAsnAspAlaAspSerThrSerValGluAlaMetTyrSerValAlaSerGlnCysLeu 437
    |||
Db 221 ATGAGCGATCGGAGCCCTGGCTGGTGAAGCGCATGTACTCTGTGCTAGCCAGTGTCTG 162
    |||

Qy 438 HisGluLysLysAsnLysArgProAspIleLysLysValGlnGlnLeuLeuGlnMet 457
    |||
Db 161 CACGAGAAAGAAAAACAGACGGCCAGACATTCGCAAGGTTCAACAGCTGTACAAAGAGACC 102
    |||

Qy 458 ThrAla 459
    |||
Db 101 TCTGCA 96
    |||

RESULT 8
BG616438 811 bp mRNA linear EST 18-APR-2001
LOCUS 602642772F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4773760 5',
DEFINITION
    mRNA sequence.
ACCESSION
    BG616438
VERSION
    BG616438.1 GI:13667809
KEYWORDS
    EST.
SOURCE
    Homo sapiens (human)
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
    1 (bases 1 to 811)
    NIH-MGC http://mgs.nci.nih.gov/.
  
```

TITLE
 JOURNAL
 COMMENT
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LICM1645 row: p column: 17
 High quality sequence start: 3
 High quality sequence stop: 613.
 Location/Qualifiers

FEATURES
 source

```

1..811
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:4773760"
    /tissue_type="embryonal carcinoma"
    /lab_host="DH10B (T1 Phage-resistant)"
    /clone_1b="NIH_MGC 61"
    /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
    SfiI (ggcgccctcgcc); Site_2: SfiI (ggcgatcattgccc);
    Double-stranded cDNA was prepared from cell line RNA. 5'
    and 3' adaptors were used in cloning as follows: 5'
    adaptor sequence: 5'-CACGCCATTTAGGC-3' and 3' adaptor
    sequence: 5'-ATTCTAGAGCGGAGCGCGCCGACATG-dT(30)BN-3'
    (where B = A, C, or G and N = A, C, G, or T). Average
    insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
    contained inserts by PCR. This library was enriched for
    full-length clones and was constructed by Clontech
    Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
    Library."
  
```

BASE COUNT 246 a 196 c 164 g 205 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.75e-97 Length: 811
 Score: 935.50 Matches: 198
 Percent Similarity: 81.10% Conservative: 8
 Best Local Similarity: 77.95% Mismatches: 32
 Query Match: 39.21% Indels: 16
 DB: 10 Gaps: 4

US-10-001-254-16 (1-460) x BG616438 (1-811)

```

Qy 1 MetAsnLysProIleThrProSerThrTyrValArgCysLeuAsnValGlyLeuIleArg 20
    |||
Db 53 ATGAACAAACCCATTAACCATCAATCAATATGCGCTCCCAATGTGACATTAATTAG 112
    |||

Qy 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrLysLysLeuAlaValAlaIleLys 40
    |||
Db 113 AAGCTCTCAGATTTTATGATCTCTCAAGAGATGGAAGAAGTACTAGCTATTAA 172
    |||

Qy 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgArgPheGluAlaLeuLeu 60
    |||
Db 173 AAACCATCTGGTGCATGATGATACATCAATTCATCAATGAAGGATTTGAAGCATTA 212
    |||

Qy 61 GlnThrGlyLysSerProThrSerGluLeuLeuPheAspTyrGlyThrThrAsnCysThr 80
    |||
Db 233 CAACTGGAANAAGTCCACTTCTGATTAATGCTGTGACGTGGGACCAACAAATTCACA 292
    |||

Qy 81 ValGlyAspLeuValAspLeuLeuIleGlnAsnGluPhePheAlaProAlaSerLeuLeu 100
    |||
Db 293 GTTGGTGAATCTGTGATCTTTTGAATCCAAATGAATTTTCGCTCTCGAGTCTTTG 352
    |||

Qy 101 LeuProAspAlaValProLysThrAlaAsnThrLeuProSerLysGluAlaIleThrVal 120
    |||
Db 353 CTCCTCAGATGCTGTCCCAAACTGTATACACTTCTTAAAGAACTTAAACAGTT 412
    |||

Qy 121 GlnGlnLysGlnMetProPheCysAspLysAspArgThrLeuMetThrProValGlnAsn 140
    |||
  
```



```

Db      413 CAGCAAAAACATGCTTCTGTGACAAAGACAGACATTGATGACACTGTGCGAGAT 472
Qy      141 LeuGluGlnSerTyrMetProPheSerSerSerProGluAsnLysSerLeuGluVal 160
Db      473 CTTGAACAAAGCATATGTCACCTGACTCTCTCAAGTCCAGAACTCAAGATTGAAAGTT 532
Qy      161 SerAspThrArgPheHisSerPheSerPheTyrGlu-LeuLysAsnValThrAsnAsn-P 180
Db      533 AGTATACACGTTTTCACAGTCTTTTCACTTTCATGATCCAGACAGATGTCAAAATTAATC 592
Qy      180 heAspGluArgProIleSerValGlyLysAsnLysMet-GlyGluGlyGlyPheGlyVal 199
Db      593 TCGATGAACAGACCACTTCTGTGCGGCTACTAAACATGGGAGAGGAGGAGATTCGCGAG 652
Qy      200 ValTyrLysGlyTyrValAsn-----AsnThrValAlaValLysLysLeuAla 216
Db      653 TTGTATCATTCACAGGCTCAATCAATACACCCCAACTGTGCGCCGCGAAGAAAGCT 712
Qy      217 Ala-----MetValAspIleThrThrGluGluLeuLysGlnGlnPhe----- 230
Db      713 TGCAGCCACCGGCTTACAT-----CACTACCGGACACACTCTGTAACAGACCT 763
Qy      231 ---AspGlnGluLysValMetAlaLysCysGln 241
Db      764 TGATCCAGAACACTAACGCACTGGCGCAAGTGTCAA 799
RESULT 9
LOCUS   BU441365 769 bp mRNA linear EST 29-NOV-2002
DEFINITION 603208981F1 CSEQRBN11 Gallus gallus cDNA clone CHEST185a21 5', mRNA
SEQUENCE
BU441365
BU441365.1 GI:25930676
EST.
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 769)
Boardman, P. E., Sanz-Ezquerro, J., Overton, I. M., Burt, D. W., Bosch, E.,
Fong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
2235534
2235534
JOURNAL MEDLINE PUBMED
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST
)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
FEATURES
Source
1..769
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer and broiler"
/db_xref="taxon:9031"
/clone="CHEST185a21"
/sex="Male and female"
/tissue_type="muscle"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQRBN11"
/note="Vector: pBluescript II KS(+); Site 1: EcoRI;
Site 2: NotI. This normalized library was constructed from
1 million independent clones. cDNA synthesis was initiated
using an oligo(dT) primer, using methylated C in the first
strand synthesis reaction. Following this first strand
reaction, double-stranded cDNA was bluntended, ligated to

```

```

BASE COUNT 230 a 146 c 180 g 213 t
ORIGIN
Alignment Scores:
Pired. No.: 3,298-90 Length: 769
Score: 874.00 Matches: 166
Percent Similarity: 87.78% Conservative: 28
Best Local Similarity: 75.11% Mismatches: 26
Query Match: 36.63% Indels: 1
DB: 13 Gaps: 0
US-10-001-254-16 (1-460) x BU441365 (1-769)
Qy      234 IleLysValMetAlaLysCysGln-HisGluAsnLeuValGluLeuGlyPheSerSe 253
Db      1 CTAGAAATTATGCAAGGTAAACAAATGATGAGATTGCTTGTTCTCAAG 60
Qy      253 rAspGlyAspAspLeuCysLeuValTyrValTyrMetProAsnGlySerLeuLeuAspAr 273
Db      61 TGATGCTGCTCAGCCCTGTTGCTGTATGAATACATGCCAATGTTGCTGCTGACAG 120
Qy      273 gLeuSerCysLeuAspGlyThrProPheSerSerTrpHisMetArgCysLysIleAlaG 293
Db      121 ACTTGCTGTGTGATGAGTCCACCTCACTTATCTTGGAACAACAAGTGAATGTGTCA 180
Qy      293 nGlyAlaAlaAsnGlyIleAsnPheLeuHisGluAsnHisIleHisArgAspIleLys 313
Db      181 AGGTACTGGGAATGCAATCACTTCTGATGACATATCATATTCACAGACATTTA 240
Qy      313 sSerAlaAsnIleLeuLeuAspGluAlaPheThrAlaLysIleSerAspPheGlyLeuAl 333
Db      241 AAGTCAATATCTTATTATACATGATACGATATGCCCCAAATTTGGACTTTGGACTTGC 300
Qy      333 aArgAlaSerGluLysPheAlaGlnThrValMetThrSerArgIleValGlyThrThrAl 353
Db      301 AAGAGCATCAGTAACATTCACAGCAACCATCATGACAGACAGATGTGGAAACAGACAG 360
Qy      353 aTyrMetAlaProGluAlaLeuArgGlyGluIleThrProLysSerAspIleTyrSerPh 373
Db      361 CTATATGGCACTGGAAGCTCTGCGAGAGAAATTAACGCTTAATCCGATATCTTCAGCTT 420
Qy      373 eGlyValValLeuLeuGluIleIleThrGlyLeuProAlaValaAspGluHisArgGluPr 393
Db      421 TGGGGTAGTCTTACTAGAAATTAACAGAGTCTGCCACAGATAGACGAAACCGGGAGGC 480
Qy      393 oGlnLeuLeuLeuAspIleLysGluGluIleGluAspGluGluLysTyrIleGluAspTy 413
Db      481 AAGATTACTGTTAAGATCAAGATCAAGAAATGAGATGAGAGGACGACATATAGAGATTAA 540
Qy      413 rIleAspLysLysMetAsnAspAlaAspSerThrSerValGluAlaMetTyrSerValAl 433
Db      541 TGTTCACGTAAAGATGAGAGTGGAGTCAACTTCAGTCAATAAATGATATCTTCACTTCC 600
Qy      433 aSerGlnCysLeuHisGlyLysLysAsnLysArgProAspIleLysLysValGlnGlnIle 453
Db      601 TGATCATGTTCTGATGAGAAAAAACAAGAGGCCAACATGAGATGTGTATGACCT 660
Qy      453 u 453
Db      661 G 661
RESULT 10
LOCUS   BF696981 719 bp mRNA linear EST 22-DEC-2000
DEFINITION 602130160F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4287014 5',

```


ACCESSION mRNA sequence.
 VERSION BF696981
 KEYWORDS BF696981.1 GI:11982389
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 719)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 REFERENCE Unpublished
 AUTHORS Contact: Robert Strausberg, Ph.D.
 TITLE Email: cga@b6-remail.nih.gov
 JOURNAL Tissue Procurement: ATCC
 COMMENT CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: http://image.llnl.gov
 Plate: LICM124 row: 9 column: 15
 High quality sequence stop: 632.
 Location/Qualifiers
 1..719
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4287014"
 /tissue_type="primitive neuroectoderm"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_1lb="NH MGC 56"
 /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctggcc); Site_2: SfiI (ggccattggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCATTATGGCC-3' and 3' adaptor sequence: 5'-ATCTAGAGCGCGCGCCGACATG-dt(30)BV-3' (where B = A, C or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 216 a 145 c 163 g 195 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3.61e-88 Length: 719
 Score: 856.00 Matches: 185
 Percent Similarity: 91.43% Conservative: 7
 Best Local Similarity: 88.10% Mismatches: 10
 Query Match: 35.88% Indels: 9
 DB: 10 Gaps: 1

US-10-001-254-16 (1-460) x BF696981 (1-719)

QY 1 MetAsnlyProIleThrProSerThyValArgCysLeuAsnValGlyLeuLea 20
 DB 68 ATGACCAACCCATCAACATCAATATGATGCTGCTCAATGTTGACCTAATTAGG 127
 QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTrpLysLysLeuAlaValAlaIleLys 40
 DB 128 AACCTGTCAGATTATGATCTCTCAAGAGCATGAGAGAGTTAGCTGATATTAA 187
 QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgArgPheGluAlaLeuLeu 60
 DB 188 AAACCACTGCTGATGATATGATACATCACTTCACTAAGAGATTGAAGCATTA 247
 QY 61 GlnThrGlyLysSerProThrSerGluLeuLeu-PheAspTrpGlyThrThrLysCys 80
 DB 248 CAAACCTGGAAGAAAGTCCCATCTTGATTAATGCTGTTGACTGGGCGCCCAATATG 307
 QY 80 rValGlyAspLeuValAspLeuLeuIleGlnAsnGluPhe-PheAlaProAlaSerLeu 100

DB 308 AGTGGGATCTGGTGGATCTTTTGGATCCAAAATGATTTGTTGGCTCCGCAAGTCTTT 367
 QY 100 euleuProAspAlaValProLysThrAlaAsnThrLeuProSerLys-GluAlaIleThr 119
 DB 368 TGGCTCCAGATGCTGTTCCCAAACTGCTAATACCTTTCAAGAGAAGCTAATA 427
 QY 120 ValGlnGlnGlnGlnMetProPheCysAspLysAspArgThr-LeuMetThrProValG 139
 DB 428 GTTCAGCAAAAACAGATGCTTCTGTGACAAAGACAGACAGTTGATGACCTGTGCA 487
 QY 139 n-AsnLeuGlnGlnSerTyrMetProProAspSer-SerSerProGlnAsnLysSerLeu 158
 DB 488 GAATCTTGAACAAGCTATATGCGCACTGCTCCAGTCCAGAACATTAACATTA 547
 QY 159 -GluValSerAspThrArgPheHisSerPheSerPheTyrGluLeuLysAsnValThr 178
 DB 548 GGAAGTTAGTGAATACAGTGTTCACAGTTTTCATTCATGAACCTGAAAGATGTCAGA 607
 QY 178 nAsnPheAspGluArgProIleSerValGlyAsnLysMetGlyGluGlyPheG 198
 DB 608 TACTTTGATGAACGACCCATTTCTGTGGTGTATTAATG3GCAC-GAAGGATTGG 666
 QY 198 yValValTyrLysGly 203
 DB 667 G---TGATATAAGCT 679

RESULT 11
 BU225249 962 bp mRNA linear EST 26-NOV-2002
 LOCUS 603947374F1 CSBQCHN23 Gallus gallus cDNA clone CHEST902c19 5', mRNA
 DEFINITION
 sequence.
 BU225249
 BU225249.1 GI:25460959
 VERSION
 KEYWORDS
 SOURCE EST.
 ORGANISM Gallus gallus (chicken)
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianidae; Gallus.
 1 (bases 1 to 962)
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 CONTACT: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES
 source
 1..962
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, Hisex"
 /db_xref="taxon:9031"
 /clone="CHEST902c19"
 /dev_stage="22"
 /lab_host="DH10B"
 /clone_1lb="CSBQCHN23"
 /note="Organ: head; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was bluntended, ligated to NotI adaptors, digested with EcoRI, site-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT 280 a 184 c 222 g 275 t 1 others

ALIGNMENT SCORES:

Alignment Scores:	5.62e-82	Length:	962
Pred. No.:	804.00	Matches:	162
Score:	68.71%	Conservative:	40
Percent Similarity:	55.10%	Mismatches:	37
Best Local Similarity:	33.70%	Indels:	56
Query Match:	13	Gaps:	4

US-10-001-254-16 (1-460) x B0225249 (1-962)

```

QY 165 PheH1SerPheSerPheTyrgIuLeuLYaSnValThraSnAnPheAepGIuaTgPro 184
DB 4 TTCACAAATTTTGGTTTCAGACTTGAAGAGTTTCAAAATATTTTGTATGACGACCA 63
QY 185 IleSerValGlyGlyAnlySmecGlyGlyGlyPheGlyValValIYrlySgLYr 204
DB 64 GAATCAGCTGGAGAAATAAGCTGGGGGAAGTGGCTTGGCAATGTGTTCAAAAGGCTAC 123
QY 205 ValAsnAnthrThrValAlaValLyLeuAlaAlaMetValAspIleThrIngu 224
DB 124 ATCAATGGGAGAAACGTTGCTGCAGAAACCTTGCTGTGTGTATGTATGTCACAG 183
QY 225 GluLeuLySgIngnPheAepGIngnIuIeLYaValMetAlaLYsCysGIngnSgLYaSn 244
DB 184 GATTGAAACAGAGCTTGTATCAAGAAATGAAATTTGTCAGAAAGTAAACATGAAAT 243
QY 245 LeuValGluLeuLySgIngnPheSerSerSgLYaAspIleuLYaValIYrValTYr 264
DB 244 CTGGTAGAATTTGCTGTGTCTCAAGTAGTGTGCTCAGCCCTGTTGGTATGATAC 303
QY 265 MetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAspGlyThrProPoluSer 284
DB 304 ATGCCCAATGCTTGGCTGTGACAGACTTGTCTGTATGATGACCTCACTATTTCT 363
QY 285 TyrH1SmecArgCysLYsIleAlaGIngnLYaAlaAnSgLYIleAsnPhelEuH1SgLY 304
DB 364 TGAACACCAAGCTGTGAATTTGCTCAAGTACGCAATGCGATCACTTCTGCAATGAC 423
QY 305 AsnH1SgIleH1SargAspIleLYsSerAlaAnIleLeuLeuAspGluAlaPheThr 324
DB 424 AATAATCATATTCACAGACATTAANAAGC----- 453
QY 325 AlaLYsIleSerAspPheGlyLeuAlaArgAlaSerGlyLYsPheAlaGInThraMet 344
DB 454 -----ATCAGG 459
QY 345 ThrSerArgIleValGlyThrThraIaTyMetAlaProGluAlaLeuArgGlyIuIe 364
DB 460 ACAAGCGTATGCTGTGAAGATGCTTTTCTTACCTTACGCTGCTC----- 507
QY 365 ThrProLYsSerAspIleTYrSerPheGlyValValLeuLeuGluIleIleThngLYe 384
DB 508 -----TACTCTTCACTGCTTGAATATATC-TTATTAACCT----- 542
QY 385 ProAlaValAspGluH1SargGluProGInLeuLeuAspIleLYsGluGluIleGlu 404
DB 543 -----GATTCGTAT-----GAATTTGAG 560
QY 405 AspGluGluLYsThrIleGluAspTYrIleAspLYsMetAsnAspAlaAspSerThr 424
DB 561 GATGAGAGCGCATATAGAGATTTATGACTTAAGATGATGATGCTGGATGCACT 620
QY 425 SerValGluAlaMetTYrSerValAlaSerGInCysLeuH1SgLYsLYsAsnLYsArg 444
DB ----- 444

```

```

DB 621 TCAGTTCATMAAATGATTCCTGCTGATCATGCTTGAATGACAAAAAACHGAGG 680
QY 445 ProAspIleLYsLYsValGIngnLeuLeuGIngnMet 457
DB 681 CCAACATCAAGAAATGCTCCACAGTATTTTACAAAGATTA 720

```

RESULT 12

LOCUS	541 bp	mRNA	linear	EST 27-MAR-2003
DEFINITION	340084 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.			
ACCESSION	BC691069			
VERSION	BC691069.1	GI:13932889		
KEYWORDS	EST.			
SOURCE	Bos taurus (cow)			
ORGANISM	Bos taurus			

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PubMed

COMMENT

USDA, ARS, Beltsville Agricultural Research Center
Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
Tel: 301 504 8416
Fax: 301 504 8414
Email: tad@lpsl.barc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -mnscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGCAACACGCTATGACCAT
BACKWARD: GTTTCACGATCAGCAGC
Plate: 100 row: B column: 15
Seq primer: ATTAGTGCACCTATGAC.
Location/Qualifiers

FEATURES

source

1..541

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/rfeature_type="pooled"

/lab_host="DH10B"

/clone_lib="BARC 5BOV"

/note="Vector: PCMV SPORT6; Site_1: NotI; Site_2: SalI; Library made from pooled mRNA isolated from mammary tissues at eight physiological, developmental, and disease states."

BASE COUNT

ORIGIN

169 a 96 c 119 g 157 t

ALIGNMENT SCORES:

Alignment Scores:	5.41e-79	Length:	541
Pred. No.:	774.50	Matches:	150
Score:	88.33%	Conservative:	9
Percent Similarity:	83.33%	Mismatches:	19
Best Local Similarity:	32.46%	Indels:	2
Query Match:	10	Gaps:	1

US-10-001-254-16 (1-460) x BG691069 (1-541)

```

QY 26 IleAspProGIngnIuLYrTyLYsLYeAlaValAlaIleLYsLYsProSergLYaSp 45
DB 2 ATGATCTCTCAAGAAAGATGGAAGAGTTCAGCTGCTATTAATAAACATCTGGTAT 61
QY 46 AspArgTYrAsnGIngnPheH1SgIleArgArgPheGIngnAlaLeuLeuGIngnThrgLYsSer 65
DB 62 GATGATATCAATCAGTTCACATTAAGAGATTTGAAGCATTTCTGCAAAATTTGAAAAAGC 121

```

OY		66	ProThrsSerGluLeuLeuPheAspTrpGlyThrTranCysThrValAlaAspleuVal	85
Db		122	CCCAAGTGTGAAGTTACTCTGTTTCACCTGGGGCACCAAAATTCACAGCTTGATCTTGTG	181
OY		86	AspleuLeuIleGlnAsnGluPhePheAlaProAlaSerLeuLeuProAspAlaVal	105
Db		182	GATATTGTGGTCCAAAATGAGTTTTTGGCCCCCTGCAGAATCTTTTCTCCACAGATCTGTA	241
OY		106	ProLysThrAlaAsnThrLeuProSerLys---GluAlaIleThrValGlnGlnLysGln	124
Db		242	CCCAAAAAATGTTAATAACACTGCTCTTCAAAGTCACTGTAAGCAGCTTCACAGAAACCG	301
OY		125	MetProPheCysAspLyAspArgThrLeuMetThrProValGlnAsnLeuGlUGlnser	144
Db		302	AAGCCCTCTGTGGCAAGAACGACACACTGTGTGATATCATGATGAGAAATCTCGAACAAAC	361
OY		145	TyrMetProProAspSerSerSerProGluAsnLysSerLeuGluValSerAspThrArg	164
Db		362	TATGTCATACCCTGACTCTCTCAAGTCCACAAAATAACAAGTTTAGAATTAGATACACGT	421
OY		165	PheHis-SerPheSerPheTyrrGluLeuLysAsnValThrAsnSnpPheAspGUARGPr	184
Db		422	TTTCAACAAGTTTTTATTTTTTGAATTAGAAGGATGTCCAATATVACITTYTATGAAACGCC	481
OY		184	oileSerValGlyGlyAsnLysMetGlyGlyUGlyValGlyPheGlyValValTyrLysGly	203
Db		482	CATTTCCTTGTGCTGTATCAAAAATGGAGAAAGAGGGTTTGGAGTTGTGTATAAGGC	539
RESULT 13				
Bb613167				
LOCUS		676 bp	mRNA	linear EST 26-OCT-2001
DEFINITION		Bb613167 RIKEN full-length enriched, 10 day neonate skin Mus		
VERSION		Bb613167	musculus cDNA clone 4732460109 5', mRNA sequence.	
KEYWORDS		Bb613167.1 GI:16453871		
SOURCE		EST.		
ORGANISM		Mus musculus (house mouse)		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.		
TITLE		1 (bases 1 to 676)		
JOURNAL		Arkawa,T., Carrinci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,		
COMMENT		Himoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,Y., Konno,H., Kouda		
		M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,		
		Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki		
		,D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,		
		Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,		
		Muramatsu,M., and Hayashizaki,Y.		
		RIKEN Mouse ESTs (Arakawa,T., et al. 2001)		
		Unpublished		
		Contact: Yoshihide Hayashizaki		
		Laboratory for Genome Exploration Research Group, RIKEN Genomic		
		Sciences Center(GSC), Yokohama Institute		
		The Institute of Physical and Chemical Research (RIKEN)		
		1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan		
		Tel: 81-45-503-9222		
		Fax: 81-45-503-9216		
		Email: genome-res@sc.riken.go.jp,		
		URL:http://genome.gsc.riken.go.jp/		
		Carrinci,P., Shibata,Y., Hayata,N., Sugahara,Y., Shibata,K., Itoh		
		,M., Konno,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.		
		Normalization and subtraction of cap-trapper-selected cDNAs to		
		prepare full-length cDNA libraries for rapid discovery of new		
		genes. Genome Res. 10 (10), 1617-1630 (2000)		
		wagui.K., Fujiwake,S., Inoue,K., Togawa,Y., Itawa,M., Ohara,E.,		
		Wachiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura		
		,S., Kawai,Y., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and		
		Hayashizaki,Y.		
		RIKEN integrated sequence analysis (RISA) system--84-Cornat		
		sequencing pipeline with 384 multicapillary sequencer. Genome Res.		
		10 (11), 1757-1771 (2000)		
		Komno,H., Fukunishi,Y., Shibata,K., Itoh.M., Carrinci,P., Sugahara		

Y. and Hayashizaki, Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001).
Kondo, S., Shingawa, A., Saito, T., Kiyosaka, H., Yamana, I., Aizawa, K., Fukuda, S., Hara, A., Itch, M., Kawai, J., Shibata, K. and Hayashizaki, Y. Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome* 12, 673-677 (2001).
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
e mouse tissues. (<http://genome.gsc.riken.go.jp>)

FEATURES

Location/Qualifiers
1. .676

```
/organism="Mus musculus"  
/mol type="mRNA"
```

```

/strain="C57BL/6J"
/db xref="taxon:10090"

```

```
/clone="4732460I09"
```

```
/sex="mixed"  
/tissue_type="skin"
```

```
/dev_stage="10 days ne
/lab_host="DH10B"
```

```
/clone_lib="RIKEN full
skin"
```

```

/Note="Site 1: SalI; S

```

prepared and sequenced
Project of Genome Expl

Genomic Sciences Center
RIKEN, Division of Experimental Research

contributed to prepare
primed with a primer

primed with a primer
GAGAGAGAGAGGATCCAGAG

prepared by using tRNA
transcriptase and sub

cap-trapper. cDNA went to Rot = 10.0 and sub

strand cDNA was prepared
sequence [5' GAGAGACG

3'}. cDNA was cloned in

Vector: a modified pBL from Lambda FLC 1"

180 a 196 c 172 g

Order.

Index	Value	Length
1.58e-76		

Clarity:	88.04%	Match
	755.00	Cons

Similarity:	76.09%	Mism
:	31.64%	Inde

10 Gaps

54-16 (1-460) X BB613167 (1-676)

1 Metasnllysprillethrrproserthrrty

16 ATGAACAAGCCGTTGACACCATCGACATA

1. **የጥያቄው አጠቃላይ፡**

lybueaseisbpfneierabpfoogll

76 AAGCTGTCGGATTATTGATCCCTCAAGA

41 LysProSerGlyAspAspArgTyrAsnG1

36 AAGCCGTCGGCGACGACAGATACATCA

51 GlnThrGlyLysSerProThrSerGluLeu

96 CAGACCGGAGAGGCCCCACCTGTGACT

31 ValGlyAspLeuValAspLeuIleGly

|||||

Db 356 GTTGGCAGCTTGTGATCTACTGTCCAGATTGAGCTGTGTCCTCCCGCCCACTCTCTG 415

Qy 101 LeuProAspAlaValProlyserThrAlaanthLeuProserLygluAlaIleThrVal 120

Db 416 CTGCGCGAATGCGCTCCCAACCGCTCAACAAAGGCTCTCTGAAGAGGCGCAACAGTG 475

Qy 121 GlnGlnLysGlnMetProPheCysAspLyAspArgThrLeuMetThrProValGlnAan 140

Db 476 GCACAAACACACGGCGCTGTGTACGAAAGAGACAGACATCCGTATAGCTTATGCCGAG 535

Qy 141 LeuGlnLysSerTyMetProProAspSerSerSerProGluAsnLysSerLeuGluVal 160

Db 536 CTAGAAACACAGCTCCAGCCAGCCAGCTCTCAAGCCACAGCAACGAAAGTGTAAAGTCC 595

Qy 161 SerAspThrArpPheHisSerPheSerPheTyGluLeuLysAsnAlaThrAsnAnpHe 180

Db 596 AGCGACACTGCTGTCACACATCTGCTTCATGAATGAAAGACATCAACAAACATCTTC 655

Qy 181 AspGluArgPro 184

Db 656 GACGACCAACCC 667

RESULT 14
BM431425 478 bp mRNA linear EST 31-JAN-2002
LOCUS 1Du016D05 Bos taurus Duodenum #1 library Bos taurus cDNA, mRNA
DEFINITION sequence.
ACCESSION BM431425
VERSION BM431425.1 GI:18453147
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE Hansen, C., Fu, A., Meng, Y., Li, C., Okine, E., Senses, C.W., Gordon, P.M.K. and Moore, S.S.
AUTHORS Gene Expression Profiling of the Bovine Gastrointestinal Tract
TITLE Unpublished
JOURNAL Contact: Dr. Stephen Moore
COMMENT Beef Genomics Laboratory
Dept of AFNS, University of Alberta
410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
Tel: 780 493 0169
Fax: 780 493 4265
Email: stephen.moore@ualberta.ca
Insert Length: 478 Std Error: 0.00
POLYA=NO.

FEATURES
source 1..478
Location/Qualifiers
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="Smooth muscle"
/cell_type="Simple columnar epithelial"
/dev_stage="Young adult"
/lab_host="X11-BlueRF strain"
/clone_lib="Bos taurus Duodenum #1 library"
/note="Organ: Intestine/Duodenum; Vector: Uni-ZAPXR; Site 1: EcorI; Site 2: Xho I"

BASE COUNT 139 a 94 c 106 g 139 t

ORIGIN

Alignment Scores:
Pred. No.: 2.15e-76 Length: 478
Score: 752.00 Matches: 147
Percent Similarity: 98.70% Conservative: 5
Best Local Similarity: 95.45% Mismatches: 2
Query Match: 31.52% Indels: 1
DB: 12 Gaps: 0

US-10-001-254-16 (1-460) x BM431425 (1-478)

Qy 239 LysCysGlnHisGluAsnLeuValGluLeuLeuGlyPheSerSerAspGlyAspAspLeu 258

Db 2 AGCTCAACATGAAATCTAGTATGATTAATCTGTTCTCAAGATGAGGATACCTCC 61

Qy 239 CysLeuValTyValTyMetProAsnGlySerLeuLeuAspArgLeuSerCysLeuAsp 278

Db 62 TCGTAGATATATGTTTACATGCGCAATGTTTCATGTCGACAGACTGCTTGAT 121

Qy 279 GlyThrProProLeuSerThrHisMetArgCysLysIleAlaGlnAlaIleAsnGly 298

Db 122 GGTACTCCCACTCTCTTGAAACATGAGATGCAAGATTGCCAGGATGCACTTATGCC 181

Qy 299 IleAsnPheLeuHisGluAsnHisIleHisIleHisArgAspIleLysSerAlaAsnIleLeu 318

Db 182 CTCAGTTATTATCATGAAACATCATATTTATTAAGATATTAAAGCAATATCTTA 241

Qy 319 LeuAspGluAlaPheThrAlaLysIleSerAspPheGlyLeuAlaArgAlaSerGluLys 338

Db 242 TTAGATGAAGACTTACAGCCAAATATCTGACTTTGGGCTTGACAGGCTTCTAGAG 301

Qy 339 PheAlaGlnThrValMetThrSerArgIleValGlyThrThrAlaTyMetAlaProGlu 358

Db 302 TTGGCCAGACAGTATGATGACAGAAATGTGGGAAACAACGCTTATATGGACCTGAA 361

Qy 359 AlaLeuArgGlyGluIleThrProLysSerAspIleTySerPheGlyValValLeuLeu 378

Db 362 GCTTGCCAGAGGAATATACACCAATCTGACATCTACAGCTTGGTGTGTTTGCTA 421

Qy 379 GluIleIleThrGlyLeuProAlaValAspGluHisArgGlu 392

Db 422 GAATA-ATTAACCTGACTTCACGCTGTGATGATGAACACCGTGAA 462

RESULT 15
BF687921 858 bp mRNA linear EST 22-DEC-2000
LOCUS 602066996P1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:406054 5',
DEFINITION mRNA sequence.
ACCESSION BF687921
VERSION BF687921.1 GI:11973329
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cga@bbs-riemail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: DLM9502 row: p column: 23
High quality sequence stop: 564.
Location/Qualifiers
source 1..858
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:406054"
/tissue_type="Glioblastoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_57"
/note="Organ: brain; Vector: pBKR-LIB (Clontech); Site 1: SfiI (ggccatcgcc); Site 2: SfiI (ggccatcgcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5'

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 18, 2004, 01:04:18 : Search time 998.41 Seconds
(without alignments)
2417.512 Million cell updates/sec

Title: US-10-001-254-26
Perfect score: 323
Sequence: 1 MNKPITPSTYVRCINVLGR.....KKPSGDDRYNGFHRCSSON 59

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 segs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10001254/runat_16012004_152423_19714/app_query.fasta_1.1109
-DB=genbml -QFMT=fastap -SFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10001254 @CGN 1.1 5671 @runat_16012004_152423_19714 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESUBSTRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELop=6 -DELEXt=7

Database :

Genbml:*
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	323	100.0	211	6 AX431316	AX431316 Sequence
2	323	100.0	2213	6 BD155790	BD155790 Primer fo
3	323	100.0	2213	9 AK027301	AK027301 Homo sapi
4	290	89.8	833	6 AR223870	AR223870 Sequence
5	290	89.8	1383	6 AX196260	AX196260 Sequence
6	290	89.8	1383	6 AX431306	AX431306 Sequence
7	290	89.8	1383	6 AF445802	AF445802 Homo sapi
8	290	89.8	1629	6 BC013316	BC013316 Homo sapi
9	290	89.8	2817	6 AX431318	AX431318 Sequence
10	290	89.8	2817	9 AF155118	AF155118 Homo sapi
11	290	89.8	2820	9 AK000528	AK000528 Homo sapi
12	290	89.8	33033	9 AY186092	AY186092 Homo sapi
13	290	89.8	118572	9 AC093012	AC093012 Homo sapi
14	286	88.5	155868	2 AC021719	AC021719 Homo sapi
15	277	85.8	501	6 AX321132	AX321132 Sequence
16	271	83.9	1542	6 AX196262	AX196262 Sequence
17	271	83.9	1542	10 AF445803	AF445803 Mus muscu
18	271	83.9	2431	10 BC051676	BC051676 Mus muscu
19	269	83.3	264657	2 AC129390	AC129390 Rattus no
20	247	76.5	294	6 AX431296	AX431296 Sequence
21	125	38.7	1719	5 BC045381	BC045381 Dario rer
22	77.5	24.0	190968	2 AC128035	AC128035 Rattus no
23	77.5	24.0	245659	2 AC127919	AC127919 Rattus no
24	77.5	24.0	270729	2 AC133702	AC133702 Rattus no
25	74	22.9	4320	9 HSCKPSRU	X64692 H.sapiens c
26	73.5	22.8	110000	2 AL954691.1	Continuation (2 of
27	73.5	22.8	184898	10 AL683822	AL683822 Mouse DNA
28	73	22.6	187380	9 AC023946	AC023946 Homo sapi
29	71.5	22.1	2190	1 AY013286	AY013286 Bacillus
30	69.5	21.5	146812	2 AC027204	AC027204 Homo sapi
31	69.5	21.5	150855	2 AC013620	AC013620 Homo sapi
32	69.5	21.5	172914	9 CNS07ECZ	AL445383 Human chr
33	69	21.4	177257	2 AC137121	AC137121 Mus muscu
34	69	21.4	197991	5 AL929109	AL929109 Zebrafish
35	69	21.4	212293	2 AC125252	AC125252 Mus muscu
36	68.5	21.2	1835	8 AF500443	AF500443 Mespilus
37	68	21.1	5187	9 AB067497	AB067497 Homo sapi
38	68	21.1	10521	1 AB014781	AB014781 Bifidobac
39	68	21.1	122435	9 AC1355481	AC1355481 Human DNA
40	68	21.1	140628	2 AC016468	AC016468 Homo sapi
41	68	21.1	172280	2 AL330879	AL330879 Human DNA
42	68	21.1	175465	2 AL356783	AL356783 Homo sapi
43	68	21.1	349980	6 AX492787	AX492787 Sequence
44	68	21.1	349980	6 AX553954	AX553954 Sequence
45	67.5	20.9	177777	2 AC022993	AC022993 Homo sapi

RESULT 1

ALIGNMENTS

AX431316
 LOCUS AX431316 211 bp DNA linear PAT 28-JUN-2002
 DEFINITION Sequence 25 from Patent WO0240680.
 ACCESSION AX431316
 VERSION AX431316.1 GI:21656185
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Pawlowski, K., Fiorentino, L., Godzik, A., Lee, S.H., Reed, J.C., Roth, W. and Stenem-Liwen, F.
 TITLE Novel death domain proteins
 JOURNAL Patent: WO 0240680-A 25 23-MAY-2002;
 BORNHAM INST (US)
 FEATURES
 source Location/Qualifiers
 1..211
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 CDS
 1..180
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD37284.1"
 /db_xref="GI:21656186"
 /translation="MNKPIPTSTYVACIANVGLIRKLSDFIDPQGMKLAVALIKRPSG
 DRIYNOFHIRCCSQN"
 BASE COUNT 77 a 41 c 36 g 57 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4 58e-36 Length: 211
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-001-254-26 (1-59) x AX431316 (1-211)
 QY 1 MetAsnLysProIleThrProSerThrTyValaArgCysLeuAsnValGlyLeuIleArg 20
 DB 1 ATGAACAAACCCATTAACACCATCATATGCGCTGCCTCATGTGACTTAATTAGG 60
 QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPlyLysLeuAlaValAlaIleLys 40
 DB 61 AACCTGCAGATTATTATGATCCTCAAGAGATGGAAGATTAGCTGTAGCTATTAA 120
 QY 41 LysProSerGlyAspAspArgTyrAenglnPheHisIleArgCysCysSerGlnAsn 59
 DB 121 AAACCATCTGTGATGATGATACATCAATTCATTACATAGATGCTGTCCCAAAAC 177
 RESULT 2
 LOCUS BD155790 2213 bp DNA linear PAT 17-JAN-2003
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BD155790
 VERSION BD155790.1 GI:27861548
 KEYWORDS JP 2002191363-A/10633.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Oca, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
 TITLE Primer for synthesizing full-length cDNA and use thereof
 JOURNAL Patent: JP 2002191363-A 10633 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 COMMENT OS Homo sapiens (human)
 PN JP 2002191363-A/10633
 PD 09-JUL-2002

PF 28-JUL-2000 JP 200028090
 PI TOSHIO OCA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 PI SAITO
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUJI OTSUKI
 PC
 C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10, C12P21/02, C12O1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 PC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 Location/Qualifiers
 (977) . (1864) .
 FT CDS
 source Location/Qualifiers
 1..2213
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 642 a 488 c 516 g 567 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5 1e-35 Length: 2213
 Score: 323.00 Matches: 59
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-001-254-26 (1-59) x BD155790 (1-2213)
 QY 1 MetAsnLysProIleThrProSerThrTyValaArgCysLeuAsnValGlyLeuIleArg 20
 DB 751 ATGAACAAACCCATTAACACCATCATATGCGCTGCCTCATGTGACTTAATTAGG 810
 QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPlyLysLeuAlaValAlaIleLys 40
 DB 811 AACCTGCAGATTATTATGATCCTCAAGAGATGGAAGATTAGCTGTAGCTATTAA 870
 QY 41 LysProSerGlyAspAspArgTyrAenglnPheHisIleArgCysCysSerGlnAsn 59
 DB 871 AAACCATCTGTGATGATGATACATCAATTCATTACATAGATGCTGTCCCAAAAC 927
 RESULT 3
 LOCUS AK027301 2213 bp mRNA linear PRI 01-AUG-2002
 DEFINITION Homo sapiens cDNA FL141395 fis, clone HEMBA1002250, weakly similar
 to PROTEIN KINASE APK1A (EC 2.7.1.1).
 ACCESSION AK027301
 VERSION AK027301.1 GI:14041890
 KEYWORDS oligo capping; fls (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Isogai, T., Oca, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Magatsuma, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuko, Y., Ninomiya, K. and Iwayanagi, T.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE
 2 (bases 1 to 2213)
 AUTHORS Isogai, T. and Otsuki, T.
 TITLE Direct Submision
 JOURNAL Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kitarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction;

5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES

SOURCE

Location/Qualifiers
1..2213
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEMBA1003250"
/issue_type="whole embryo, mainly head"
/clone_id="HEMBA1"
/dev_stage="embryo, 10 weeks"
/note="cloning vector: PME18SFL3"

BASE COUNT 642 a 488 c 516 g 567 t

ALIGNMENT SCORES:

Pred. No.: 5.1e-35 Length: 2213
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-001-254-26 (1-59) x AK027301 (1-2213)

QY 1 MetAsnLysProlleThrProSerThrTyValArgCysLeuAnValGlyLeuIleArg 20

DB 751 ATGAACAAACCCATTAACCATCAATATATGCGCTCAATGTTGACTAATTAGG 810

QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPlyLysLeuAlaValAlaIleLys 40

DB 811 AAGCTGTAGATTATTATGATCCTCAGAGAAGATGGAAGATTAGCTTACGTTATTTAA 870

QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgCysSerGlnAsn 59

DB 871 AAACATCTGTGATGATGATACATCATGTTTACATTAAGATGCTGTTCCCAAAAC 927

RESULT 4 AR223870 833 bp DNA linear PAT 26-SEP-2002

LOCUS AR223870 Sequence 10 from patent US 6440663.

DEFINITION AR223870

ACCESSION AR223870

VERSION AR223870.1 GI:23332452

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

BASE COUNT 273 a 155 c 179 g 226 t

ORIGIN

Alignment Scores:

Pred. No.: 7.64e-31 Length: 833

Score: 290.00 Matches: 54

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 89.78% Indels: 0

DB: Gaps: 0

US-10-001-254-26 (1-59) x AR223870 (1-833)

QY 1 MetAsnLysProlleThrProSerThrTyValArgCysLeuAnValGlyLeuIleArg 20

DB 50 ATGAACAAACCCATTAACCATCAATATATGCGCTCAATGTTGACTAATTAGG 109

QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPlyLysLeuAlaValAlaIleLys 40

DB 110 AAGCTGTAGATTATTATGATCCTCAGAGAAGATGGAAGATTAGCTTACGTTATTTAA 169

QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArg 54

DB 170 AAACATCTGTGATGATGATACATCATGTTTACATTAAGG 211

RESULT 5 AX196260 1383 bp DNA linear PAT 28-AUG-2001

LOCUS AX196260 Sequence 2 from Patent WO0151641.

DEFINITION AX196260

ACCESSION AX196260

VERSION AX196260.1 GI:15386462

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

BASE COUNT 463 a 244 c 283 g 393 t

ORIGIN

Alignment Scores:

Pred. No.: 1.28e-30 Length: 1383

Score: 290.00 Matches: 54

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 89.78% Indels: 0

DB: Gaps: 0

US-10-001-254-26 (1-59) x AX196260 (1-1383)

QY 1 MetAsnLysProlleThrProSerThrTyValArgCysLeuAnValGlyLeuIleArg 20

DB 1 ATGAACAAACCCATTAACCATCAATATATGCGCTCAATGTTGACTAATTAGG 60

QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPlyLysLeuAlaValAlaIleLys 40

DB 61 AAGCTGTAGATTATTATGATCCTCAGAGAAGATGGAAGATTAGCTTACGTTATTTAA 120

QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArg 54

DB 121 AAACATCTGTGATGATGATACATCATGTTTACATTAAGG 162

RESULT 6 AX413106

LOCUS AX431306 1383 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 15 from Patent WO0240680.
ACCESSION AX431306
VERSION AX431306.1 GI:21656175
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Pawlowski, K., Fiorentino, L., Godzik, A., Lee, S. H., Reed, J. C.,
TITLE Roth, W. and Stenmer-Liwen, F.
JOURNAL Novel death domain proteins
Patent: WO 0240680-A 15 23-MAY-2002;
BURNHAM INSTR (US)
FEATURES
source Location/Qualifiers
1..1383
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..1383
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD37279.1"
/db_xref="GI:21656176"
/translation="MNKPTSTYVRCINVLIRKLSDFIDPQEGMKLAVALIKKPSG
DNRVNOFHRRERFALLQTKSPTELLPDMGTNTAGDVLDLLQNEFRFASILLP
DAVPTKANTLPSKEAITVOOKMPFCDKDTLMTVPONLEOSYMPDSSPEKSLIEV
SDTRPHSFSEYELKNVTNNFDERPISVGNNKMBEGGFGVYKGVYNNNTVAVKLAAM
VDITTEELKQCFDQEI KYMAKCOHENLVELLGFSSDGDLCIVVYMPNGSLDLRLSC
LDGTPPLSWHMRCKIAOGANGINFLHENHHRDIKSANILLDEAFPAKISDPGLAR
ASEKPAQVMTSRIVGTTAYMAPEALRGELTPKSDIYSFGVLLLEITITGLPAVDHRE
POLLIDIKREIDEKTEIDYIDKKNMDASTSVAMYSVASQCLHEKKNRPDIKKV
QQLLEMTAS"
BASE COUNT 463 a 243 c 283 g 394 t
ORIGIN
Alignment Scores:
Pred. No.: 1,28e-30 Length: 1383
Score: 290.00 Matches: 54
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.78% Indels: 0
Gaps: 0
US-10-001-254-26 (1-59) x AX431306 (1-1383)
QY 1 MetAsnLysPProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
Db 1 ATGAACAACCCATTAACACCATCAACATATGTCGCTGCCTCAATGTTGACTAATTAGG 60
QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTTPlyLysLeuAlaValAlaIleLys 40
Db 61 AAGCTGTCAATTTTATTTATGATCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 41 LysProSerGlyAspAspArgTyArgGlnGlnPheHisIleArg 54
Db 121 AAACCATCTGCTGATGATGATACATCAATCACTTTTCACTAAGG 162
RESULT 7
AF445802 1383 bp mRNA linear PRI 20-APR-2002
LOCUS Homo sapiens interleukin-1 receptor associated kinase 4 (IRAK4)
DEFINITION mRNA, complete cds.
ACCESSION AF445802
VERSION AF445802.1 GI:20219009
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Li, S., Strejlow, A., Fontana, E. J. and Wesche, H.

TITLE IRAK-4: a novel member of the IRAK family with the properties of an
IRAK-kinase
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (8), 5567-5572 (2002)
MEDLINE 21957277
PUBMED 11960013
REFERENCE 2 (bases 1 to 1383)
AUTHORS Suzuki, N., Suzuki, S., Duncan, G. S., Miller, D. G., Wada, T.,
Mitsuo, C., Takada, H., Wakeham, A., Irie, A., Li, S., Penninger, J. M.,
Wesche, H., Ohnishi, P. S., Mak, T. W. and Yeh, W. C.
TITLE Severe impairment of interleukin-1 and Toll-like receptor
signalling in mice lacking IRAK-4
JOURNAL Nature 416 (6882), 750-756 (2002)
MEDLINE 11923871
PUBMED 121959395
REFERENCE 3 (bases 1 to 1383)
AUTHORS Li, S., Strejlow, A., Fontana, E. J. and Wesche, H.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-2001) Biology I, Tularik Inc., 2 Corporate Drive,
South San Francisco, CA 94080, USA
FEATURES
source Location/Qualifiers
1..1383
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1383
/gene="IRAK4"
1..1383
/gene="IRAK4"
/note="protein kinase"
/codon_start=1
/product="interleukin-1 receptor associated kinase 4"
/protein_id="AA015772.1"
/db_xref="GI:20219010"
/translation="MNKPTSTYVRCINVLIRKLSDFIDPQEGMKLAVALIKKPSG
DNRVNOFHRRERFALLQTKSPTELLPDMGTNTAGDVLDLLQNEFRFASILLP
DAVPTKANTLPSKEAITVOOKMPFCDKDTLMTVPONLEOSYMPDSSPEKSLIEV
SDTRPHSFSEYELKNVTNNFDERPISVGNNKMBEGGFGVYKGVYNNNTVAVKLAAM
VDITTEELKQCFDQEI KYMAKCOHENLVELLGFSSDGDLCIVVYMPNGSLDLRLSC
LDGTPPLSWHMRCKIAOGANGINFLHENHHRDIKSANILLDEAFPAKISDPGLAR
ASEKPAQVMTSRIVGTTAYMAPEALRGELTPKSDIYSFGVLLLEITITGLPAVDHRE
POLLIDIKREIDEKTEIDYIDKKNMDASTSVAMYSVASQCLHEKKNRPDIKKV
QQLLEMTAS"
BASE COUNT 463 a 244 c 283 g 393 t
ORIGIN
Alignment Scores:
Pred. No.: 1,28e-30 Length: 1383
Score: 290.00 Matches: 54
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.78% Indels: 0
Gaps: 0
US-10-001-254-26 (1-59) x AF445802 (1-1383)
QY 1 MetAsnLysPProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
Db 1 ATGAACAACCCATTAACACCATCAACATATGTCGCTGCCTCAATGTTGACTAATTAGG 60
QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTTPlyLysLeuAlaValAlaIleLys 40
Db 61 AAGCTGTCAATTTTATTTATGATCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 41 LysProSerGlyAspAspArgTyArgGlnGlnPheHisIleArg 54
Db 121 AAACCATCTGCTGATGATGATACATCAATCACTTTTCACTAAGG 162
RESULT 8
BC013316 1629 bp mRNA linear PRI 04-SEP-2001
LOCUS Homo sapiens, clone MGC:13330 IMAGE:4287014, mRNA, complete cds.
ACCESSION BC013316
VERSION BC013316.1 GI:15426431

KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS 1 (bases 1 to 1629)
TITLE Strausberg, R.
JOURNAL Direct Submission
Submitted (31-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
Contact: amadan@systemsbio.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettelman and Anuradha Madan
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 19 Row: n Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7705840.
FEATURES
source
1. 1629
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:13330 IMAGE:4287014"
/tissue_type="Brain, primitive neuroectodermal"
/clone_id="NIH_MGC_56"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
71..1453
/codon_start=1
/product="unknown (protein for MGC:13330)"
/protein_id="AAH13316.1"
/db_xref="GI:15426432"
/translation="MNKPIPTSTYVRCINVGILRKISDFIDPQSGKKLAVALIKKPSG
DDRINQPHIRREFALLQTKSPSELPFMDGTTNCTGDLVDLITQNEFFAPASLLIP
DAVPKANTLPSKEALTYQKQMPFCDKRTMTVPQNLQSYMPDSSPENKSLIEV
SDTRFHSFSEYELKNVTNNPDERPI SVGNKMGEGGVYKGVNNTTVAVKLLAM
VDITTEELKQDFQDEIKYMAKCOHENLVELLGSDGDDCLVYVYMPNGSLDRISC
LDGTPPLSMHMRCKIAQGANINFLHNHHIRHDIKSANILLDEAFTRKISDQGLAR
ASERPAQTYMTSRTVGTTYVMAPEALRGITTPKSDIYKGVILLEITIGLPAYDERRE
POLLIDIKSEIEDEKTIEDYIDKKNADSDTSVEAMYSAGSQCRHEKKNKSPDIKIV
QQLQENTAS"
BASE COUNT 542 a 291 c 328 g 468 t
ORIGIN
Alignment Scores:
Pred. No.: 1.52e-30 Length: 1629
Score: 290.00 Matches: 54
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.78% Indels: 0
Gaps: 0
US-10-001-254-26 (1-59) x BC013316 (1-1629)
QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
DB 71 ATGAACAACCCATACCATCAACATATGTGGCTGCCTCATGTGACTAATAGG 130
QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrLysLysLeuAlaValAlaIleLys 40
DB 131 AAGCTGCAGATTATATGATCTCTCAAGAAGATGAGAAAGATTAGCTGATATTA 190

QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArg 54
DB 191 AACCATCTGTGATGATGATACATCAATCACTTACATTAAGG 232
RESULT 9
LOCUS AX431318 2817 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 27 from Patent WO0240680.
ACCESSION AX431318
VERSION AX431318.1 GI:21656187
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS 1
1 Pawlowski, K., Fiorentino, L., Godzik, A., Lee, S.H., Reed, J.C.,
Roch, W. and Stenner-Jensen, F.
TITLE Novel death domain proteins
JOURNAL Patent: WO 0240680-A 27 23-MAY-2002;
BURNHAM INST (US)
FEATURES
source
1. 2817
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
50..1432
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD37285.1"
/db_xref="GI:21656188"
/translation="MNKPIPTSTYVRCINVGILRKISDFIDPQSGKKLAVALIKKPSG
DDRINQPHIRREFALLQTKSPSELPFMDGTTNCTGDLVDLITQNEFFAPASLLIP
DAVPKANTLPSKEALTYQKQMPFCDKRTMTVPQNLQSYMPDSSPENKSLIEV
SDTRFHSFSEYELKNVTNNPDERPI SVGNKMGEGGVYKGVNNTTVAVKLLAM
VDITTEELKQDFQDEIKYMAKCOHENLVELLGSDGDDCLVYVYMPNGSLDRISC
LDGTPPLSMHMRCKIAQGANINFLHNHHIRHDIKSANILLDEAFTRKISDQGLAR
ASERPAQTYMTSRTVGTTYVMAPEALRGITTPKSDIYKGVILLEITIGLPAYDERRE
POLLIDIKSEIEDEKTIEDYIDKKNADSDTSVEAMYSAGSQCRHEKKNKSPDIKIV
HOLLQENTAS"
BASE COUNT 912 a 547 c 586 g 772 t
ORIGIN
Alignment Scores:
Pred. No.: 2.66e-30 Length: 2817
Score: 290.00 Matches: 54
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.78% Indels: 0
Gaps: 0
US-10-001-254-26 (1-59) x AX431318 (1-2817)
QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
DB 50 ATGAACAACCCATACCATCAACATATGTGGCTGCCTCATGTGACTAATAGG 109
QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrLysLysLeuAlaValAlaIleLys 40
DB 110 AAGCTGCAGATTATATGATCTCTCAAGAAGATGAGAAAGATTAGCTGATATTA 169
QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArg 54
DB 170 AACCATCTGTGATGATGATACATCACTTACATTAAGG 211
RESULT 10
LOCUS AF155118 2817 bp mRNA linear PRI 05-JAN-2000
DEFINITION Homo sapiens putative protein kinase NY-REN-64 antigen mRNA,
complete cds.
ACCESSION AF155118
VERSION AF155118.1 GI:5360130

KEYWORDS	
SOURCE	Homio sapiens (human)
ORGANISM	Homio sapiens
REFERENCE	1 (bases 1 to 2817) Scalan,M.J., Gordon,J.D., Williamson,B., Stockert,E., Bander,N.H., Jongeneel,V., Gure,A.O., Jager,D., Jager,E., Knuth,A., Chen,Y.-T. and Old,L.J. Antigens recognized by autologous antibody in patients with renal-cell carcinoma Int. J. Cancer 83 (4), 456-464 (1999)
AUTHORS	2 (bases 1 to 2817) Scalan,M.J., Gordon,J.D., Williamson,B., Stockert,E., Bander,N.H., Jongeneel,V., Gure,A.O., Jager,D., Jager,E., Knuth,A., Chen,Y.-T. and Old,L.J.
TITLE	Direct Submission
JOURNAL	Submitted (28-MAY-1999) Ludwig Institute, Sloan-Kettering Institute, 1275 York Ave, New York, NY 10021, USA
REFERENCE	Location/Qualifiers
AUTHORS	1..2817
FEATURES	source
source	/organism="Homio sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /cell_line="1973/10-4" /cell_type="renal cell carcinoma" 50..1432 /codon_start=1 /product="putative protein kinase NY-REN-64 antigen" /protein_id="A042884.1" /db_xref="GI:5360131" /translation="MNKPTSTYVRCINVLIRKLSDFIDPQSGMKLVAIKKPGSGDVRNPFIRIRFPAALQTKGKPTSELFDWCTTCVGDVLDLIONEFPAPSLILPDAVRKATINISKEALIVQKGFPCDRIPLMVEQSYMPDSSSPKSLFVSDTPHSEFVELKNVTMFERPIISGVMKGSGGVYKGVNNTVVAVKLLAAMDITFHSBELKQGPDEIKKMACQCHENVYELGSPSSDGDUCIVYVMPNGLDRLSLC LDTGPPSLMWRCKIAGGAAGINFLHNHHRDITSANTLLDEATAKISDFGLARASEKFAQTVMSTRIVGTAYVAPALREIIPKSDISFGVLLIITGCPADVDEHREPEQLLDIDKEIEDEKTEIDYIDKKMDADSTVEAMYSAGSQCRHEKKNKSPDKIKVHLLQEMTAS"
BASE COUNT	912 a 547 c 586 g 772 t
ORIGIN	
Alignment Scores:	
Pred. No.:	2,66e-30 Length: 2817
Score:	290.00 Matches: 54
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatch: 0
Query Match:	89.78% Indels: 0
DB:	9 Gaps: 0
US-10-001-254-26 (1-59) x AF155118 (1-2817)	
Oy	1 MetanlySPProileThrProSerThrTyValaIrgCySleuAenValIGlyleuIlearg 20
Db	50 ATGAACAACCAATACACATCAATATGATGCGCTCCTCAATGTTGACATAATTAGG 109
Oy	21 LysleuSerAaPhaIleaspProGlnGluGlyTTPpIlySlySleuAlaValAlaIleIlys 40
Db	110 AAGCTGTGAGATTATTAATGATCTCTCAAGAAAGATGGAAGATTGACGTGCTAATTAAA 169
Oy	41 LysProSerGlyAaSPaSPaRgTyraGlnPheHisIlearg 54
Db	170 AAACCATCTGGATGATATGATACATCAATCAAGTTTCACATTAAGG 211
RESULT 11	
LOCUS	AK000528 2820 bp mRNA linear PRI 22-FEB-2000
DEFINITION	Homio sapiens cDNA FLJ20521 fis, clone KAT10395.
ACCESSION	AK000528.1 GI:7020683
VERSION	

oligo capping; fls (full insert sequence).
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (sites)
 Wetanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
 Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
 Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2820)
 Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
 Shibahara,T., Tanaka,T. and Nakamura,Y.
 Direct Submission
 Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Depent of Virology, Shirokane-dai, 4-6-1,
 Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp,
 Tel:81-3-5449-5586, Fax:81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan, cDNA full insert
 sequencing: Research Association for Biotechnology; cDNA library
 construction, 5'- & 3'-end one pass sequencing: Department of
 Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).
 FEATURES
 source
 1. 2820
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="KAT10395"
 /cell_line="KATO III"
 /cell_type="signal-ring cell carcinoma"
 /clone_1b="KAT"
 /note="cloning vector pME18SFL3"
 30. 1412
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAA91232.1"
 /db_xref="GI:7020684"
 /translation="MNKPIPTSTYVRCINLVGLIRKLSDFIDPQGMKKLVAIIKKPSG
 DRRNOPIRFRFALLIOTGKSPISSELLFDGNTKICVDVLLIIONEFAPABLLRP
 DAYPKANTIPSKKAITVOCKMPCPKDPTNPNIONEIOSYMPDPSSPENSLELV
 SDTRFHSFTELEKNVTNNRDERPIISGKNMKGEGGVYKGYNNTPYAVKGLAM
 VDIITEELKQDFQDEIKVMKQCHENLVLLGSSDDDCLVIVVYVNPNSLDRLSC
 LDGPPILSWMKRCIKIAGAGNINFLNHNHRIHIDISANILLDEAPTAKISDGLAR
 ASEKFOATWTSIRVGTGAYMAPALFAGLEITPSCDIYSFVLLLEITIGPAVDEHEE
 POLLIDKEIKIEDEBKTEIDYIDKKNMDADSTSEAVYSVASQCLHEKKNRPRDKKY
 QQLQLENTAS"
 BASE COUNT 940 a 534 c 579 g 767 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.67e-30 Length: 2820
 Score: 290.00 Matches: 54
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 89.78% Indels: 0
 DB: 9 Gaps: 0
 US-10-001-254-26 (1-59) x AK000528 (1-2820)
 QY 1 MetAsnlySProlieThrProserThyTyValArgCysleuAenValGlyleuileArg 20
 Db 30 ATGAACAACCAACCAACCAACCAACCAATATGCGCTGCCTCATGTGGACTAATTAAG 89
 QY 21 LysleuSerAepPheileaspProGlnGlnGlyTrpIySylsleuAlaValAlaIlelys 40
 Db 90 AAGCTGCACATTTATTATGATCCTCAAGAAGATGGAAGAAGTACGTGTGCTATTAAA 149
 QY 41 LysProSerGlyAaspAapArgTyTAsnGlnPheIstlearg 54

Db 150 AACCATCTGGTGGATGATAGATACATCATGTTTCACATAGG 191

RESULT 12

LOCUS AY186092 33033 bp DNA linear PRI 02-DEC-2002

DEFINITION Homo sapiens interleukin-1 receptor-associated kinase 4 (IRAK4)

ACCESSION AY186092

VERSION AY186092

KEYWORDS gene, complete cds.

SOURCE AY186092.1 GI:2600792

ORGANISM Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 33033)

AUTHORS Rieder,M.J., Arnel,T.Z., Carrington,D.P., Ozuna,M., Kulaneek,S.A., Rajkumar,N., Toth,E.J., Yi,Q. and Nickerson,D.A.

TITLE Direct Submission

JOURNAL Submitted (22-NOV-2002) Genome Sciences, University of Washington, 1705 NE Pacific, Seattle, WA 98195, USA

COMMENT To cite this work please use: SeattleSNPs, NHLBI HL66682 Program for Genomic Applications, UW-PHRC, Seattle, WA (URL: <http://pga.gs.washington.edu>).

FEATURES

source 1..33033

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

34..88

/rpt_family="L2"

/rpt_type=dispersed

89..383

/rpt_family="Alu"

/rpt_type=dispersed

92..94

/frequency="0.02"

/replace="a"

133

/frequency="0.06"

/replace="a"

384..599

/rpt_family="L2"

/rpt_type=dispersed

642

/frequency="0.02"

/replace="c"

708..941

/rpt_family="L2"

/rpt_type=dispersed

854

/frequency="0.02"

/replace="a"

1265

/frequency="0.01"

/replace="c"

1533..1534

/frequency="0.93"

/replace="a"

1894..31033

/gene="IRAK4"

join(1894..1933,11040..11209,14157..14302,15117..15299,15849..16009,16902..16966,20567..20681,21130..21239,25244..25427,26599..26661,29336..29494,29617..31033)

/gene="IRAK4"

/product="interleukin-1 receptor-associated kinase 4"

2315

/gene="IRAK4"

/frequency="0.01"

/replace="c"

2365

/gene="IRAK4"

/frequency="0.01"

/replace="c"

variation

2456

/gene="IRAK4"

/frequency="0.01"

/replace="t"

2506..2807

/rpt_family="Alu"

/rpt_type=dispersed

2940

/gene="IRAK4"

/frequency="0.04"

/replace="t"

3023

/gene="IRAK4"

/frequency="0.09"

/replace="t"

3025..3915

/rpt_family="L2"

/rpt_type=dispersed

3038

/gene="IRAK4"

/frequency="0.01"

/replace="c"

3115..3129

/gene="IRAK4"

/frequency="0.72"

/replace="a"

3189

/gene="IRAK4"

/frequency="0.01"

/replace="g"

3536

/gene="IRAK4"

/frequency="0.14"

/replace="c"

3573

/gene="IRAK4"

/frequency="0.01"

/replace="t"

3819

/gene="IRAK4"

/frequency="0.17"

/replace="t"

3877

/gene="IRAK4"

/frequency="0.01"

/replace="g"

4144

/gene="IRAK4"

/frequency="0.04"

/replace="a"

4438..4610

/rpt_family="MIR"

/rpt_type=dispersed

4602

/gene="IRAK4"

/frequency="0.01"

/replace="t"

4797

/gene="IRAK4"

/frequency="0.02"

/replace="a"

4972

/gene="IRAK4"

/frequency="0.31"

/replace="a"

5759..5797

/rpt_family="L2"

/rpt_type=dispersed

5821

/gene="IRAK4"

/frequency="0.02"

/replace="g"

5880

/gene="IRAK4"
 /frequency="0.01"
 /replace="c"
 5554
 /gene="IRAK4"
 /frequency="0.01"
 /replace="c"
 5564 .6036
 /rpt_family="MERL_type"
 /rpt_type=dispersed
 6023 .6590
 /rpt_family="MERL_type"
 /rpt_type=dispersed
 6204
 /gene="IRAK4"
 /frequency="0.04"
 /replace="c"
 6661 .6662
 /gene="IRAK4"
 /frequency="0.04"
 /replace="c"
 6998
 /gene="IRAK4"
 /frequency="0.04"
 /replace="g"
 7015 .7073
 /rpt_family="Mariner"
 /rpt_type=dispersed
 7074 .7366
 /rpt_family="Alu"
 /rpt_type=dispersed
 7122
 /gene="IRAK4"
 /frequency="0.01"
 /replace="a"
 7147
 /gene="IRAK4"
 /frequency="0.01"
 /replace="a"
 7234
 /gene="IRAK4"
 /frequency="0.04"
 /replace="c"
 7367 .8636
 /rpt_family="Mariner"
 /rpt_type=dispersed
 7392
 /gene="IRAK4"
 /frequency="0.01"
 /replace="c"
 7597
 /gene="IRAK4"
 /frequency="0.01"
 /replace="c"
 7794
 /gene="IRAK4"
 /frequency="0.09"
 /replace="g"

Alignment Scores:
 Pred. No.: 3.32e-29 Length: 33033
 Score: 290.00 Matches: 54
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 89.78% Indels: 0
 DB: 9 Gaps: 0

US-10-001-254-26 (1-59) x AY186092 (1-33033)
 QY 1 MetAsnlySProilethrProSerThrTyValAlrGcysleuSantValGlyLeuIleAlrG 20
 DB 11049 ATGAACAAACCCATTAACACCATCATATGTGGCTGCTCAATGTGACATTAATAGG 11108

QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPylsYleuAlaValAlaIleLys 40
 DB 11109 AAGCTGTAGATTATTATTCCTCCTAAGAGAGTGAAGAGTGAAGCTGTAGCTATTATAA 11168
 QY 41 LysProSerGlyAspAspArgTyrTranGlnPheHisIleAlrG 54
 DB 11169 AAACCATCTGTGATGATGATACATCAATCATGTTCACTAATAGG 11210

RESULT 13
 AC093012
 LOCUS AC093012
 DEFINITION Human BAC library complete sequence.
 ACCESSION AC093012
 VERSION AC093012
 KEYWORDS HTG.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaratunga, H.C., Are, J.R., Ayele, M., Banks, T., Barbarta, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chaver, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.B., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulik, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Kovach, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L.J., Li, J., Li, Z., Licharge, O., Lien, C., Liu, J., Liu, W., Louised, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metscher, S., Metzker, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Ollis, M., Ren, Y., Rives, M., Rojas, A., RojudoKan, I., Rolfe, W., Ruiz, S., Severy, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shoshitari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabors, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umami, K., Vaequez, L., Vera, V., Villalón, D., Vinson, R., Wang, O., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Kuchertapali, R., Weinstein, G., and Gibbs, R.

TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 118572)
 AUTHORS Worley, K.C.
 JOURNAL Direct Submission
 Submitted (09-AUG-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 118572)
 AUTHORS Worley, K.C.

TITLE Direct Submission
JOURNAL Submitted (04-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
4 (bases 1 to 118572)
Direct Submission
Submitted (21-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
5 (bases 1 to 118572)
Direct Submission
Submitted (29-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
6 (bases 1 to 118572)
Direct Submission
Submitted (03-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
7 (bases 1 to 118572)
Direct Submission
Submitted (24-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT
On Jun 29, 2002 this sequence version replaced gi:21535906.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

<http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES

source
1. 118572
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-210N13"

misc_feature
complement(1..2004)
/note="overlaps bases 1..2004 of clone AC016143"
/function="clone overlap"

repeat_region
869..903
/rpt_family="HSMAR1"

repeat_region
1047..1467
/rpt_family="MLTK"

repeat_region
1644..1760
/rpt_family="MIR"

repeat_region
1985..2295
/rpt_family="AluY"

repeat_region
2296..2325
/rpt_family="AT_rich"

repeat_region
complement(2730..3024)
/rpt_family="AluSg"

repeat_region
4082..4211
/rpt_family="AluSg"

repeat_region
4389..4415
/rpt_family="MERSA"

repeat_region
4502..4561
/rpt_family="AT_rich"

repeat_region
complement(4612..4778)
/rpt_family="MIR"

repeat_region
5859..5926
/rpt_family="AT_rich"

repeat_region
complement(6385..6636)
/rpt_family="AluSg"

repeat_region
complement(7252..7562)
/rpt_family="L2"

repeat_region
7574..8066
/rpt_family="MLTIF"

repeat_region
complement(8265..8438)
/rpt_family="MIR"

repeat_region
8889..8927
/rpt_family="AT_rich"

repeat_region
10113..10252
/rpt_family="L1M4"

repeat_region
10433..10726
/rpt_family="AluY"

repeat_region
11181..11552
/rpt_family="THB1C"

repeat_region
complement(11739..11889)
/rpt_family="MERSB"

repeat_region
11910..12124
/rpt_family="MERSA"

repeat_region
12249..12617
/rpt_family="THB1C"

repeat_region
14290..14310
/rpt_family="AT_rich"

repeat_region
14363..14670
/rpt_family="AluSg"

repeat_region
complement(14742..15039)
/rpt_family="AluSx"

repeat_region
15669..15691
/rpt_family="TTGn"

repeat_region
complement(15692..15972)
/rpt_family="AluSc"

repeat_region
16265..16352
/rpt_family="L1PAL6"

repeat_region
16353..16631
/rpt_family="AluJo"

repeat_region
16632..17007
/rpt_family="L1PAL6"

repeat_region
17329..17356

Alignment Scores:

1. 23e-28
Pred. No.: 118572
Score: 290.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 89.78%
DB: 9
Length: 118572
Matches: 54
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

```

US-10-001-254-26 (1-59) x AC093012 (1-118572)
QY 1 MetasulphoproteinTherProserThyYrValArgCysLeuAsnValGlyLeuIleArg 20
Db 35063 ATGACAAACCCATACACCATACATATGCGCTGCTCAATTTGGACTAATTAAG 35122
QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrIleLysLeuAlaValAlaIleLys 40
Db 35123 AAGCTGTCAAGATTTTATTGATCTCTCAAGAAGATGAGAGAGATTACTGACTATTTAA 35182
QY 41 LysProSerGlyAspAspArgTyrArgGlnPheHisIleArg 54
Db 35183 AAACCATCTGGTATGATAGATACATCACTGTTTCACATTAAG 35224
RESULT 14
AC021719/c 165868 bp DNA linear HNG 24-AUG-2002
LOCUS Homo sapiens chromosome 19 clone RP11-210N13 map 19, WORKING DRAFT
DEFINITION SEQUENCE, 14 unordered pieces.
ACCESSION AC021719
VERSION AC021719.3 GI:7408021
KEYWORDS HNG; HNGS PHASE1; HNGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 165868)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Homo sapiens chromosome 19, clone RP11-210N13
REFERENCE 2 (bases 1 to 165868)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A.,
Choepe,I., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Dominko,M., Doyle,M., Fenebor,J.,
Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatae,A., Klein,J.,
Lander,T., Lehoccky,J., Levine,R., Liu,C., Liu,G., Locke,K.,
Macedonald,P., Margulis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
Pierre,N., Pisanu,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talame,J., Testfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (19-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 165868)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepe,I., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodgson,S., Dominko,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Coyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatae,A.,
Klein,J., LaRoque,K., Lamazeres,R., Landers,C., Kann,L., Karatae,A.,
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisanu,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talame,J.,
Testfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

```

```

TITLE Direct Submission
JOURNAL Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
COMMENT Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 4, 2000 this sequence version replaced gi:6957781.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5945
Center clone name: 210 N 13
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 156294 bases at least Q40
Consensus quality: 160271 bases at least Q30
Consensus quality: 162203 bases at least Q20
Insert size: 175000; agarose-fp
Insert size: 164568; sum-of-contigs
Quality coverage: 4.9 in Q20 bases; agarose-fp
Quality coverage: 5.2 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1982: contig of 1982 bp in length
1 1983 2082: gap of 100 bp
2083 4169: contig of 2087 bp in length
4170 4269: gap of 100 bp
4270 8449: contig of 4180 bp in length
8450 8549: gap of 100 bp
8550 15400: contig of 6851 bp in length
15401 15500: gap of 100 bp
15501 20885: contig of 5385 bp in length
20886 20985: gap of 100 bp
20986 27150: contig of 6165 bp in length
27151 27250: gap of 100 bp
27251 36127: contig of 8877 bp in length
36128 36227: gap of 100 bp
36228 47299: contig of 11072 bp in length
47300 47399: gap of 100 bp
47400 63394: contig of 15995 bp in length
63395 63494: gap of 100 bp
63495 78310: contig of 14816 bp in length
78311 78410: gap of 100 bp
78411 93139: contig of 14729 bp in length
93140 93239: gap of 100 bp
93240 107624: contig of 14385 bp in length
107625 107724: gap of 100 bp
107725 123977: contig of 16253 bp in length
123978 124077: gap of 100 bp
124078 165868: contig of 41791 bp in length.
Location/Qualifiers
1. 165868
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19"
/clone="RP11-210N13"
/clone_lib="RPCT-11 Human Male BAC"
1. 1982
/note="assembly_fragment"
FEATURES
source
misc_feature

```



```
misc_feature 2083..4169
/note="assembly_fragment"
clone_end:"7"
vector_side:right"
misc_feature 4270..8449
/note="assembly_fragment"
8550..15400
/note="assembly_fragment"
15501..20885
/note="assembly_fragment"
20986..27150
/note="assembly_fragment"
27251..36127
/note="assembly_fragment"
36228..47299
/note="assembly_fragment"
47400..63394
/note="assembly_fragment"
63495..78310
/note="assembly_fragment"
78411..93139
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
misc_feature 93240..107624
/note="assembly_fragment"
107725..123977
/note="assembly_fragment"
124078..165868
/note="assembly_fragment"
BASE COUNT 50117 a 31341 c 32315 g 50785 t 1310 others
ORIGIN
```

```
Alignment Scores:
Pred. No.: 6.29e-28 Length: 165868
Score: 286.00 Matches: 53
Percent Similarity: 98.15% Conservative: 0
Best Local Similarity: 98.15% Mismatches: 1
Query Match: 88.54% Indels: 0
DB: 2 Gaps: 0
```

US-10-001-254-26 (1-59) x AC021719 (1-165868)

```
QY 1 MetAsnlySProlleThrProSerThrYrValArgCysLeuAsnValGlyLeuIleArg 20
|||||
Db 114811 ATGACAAACCCATACACATCATATCTGCGCTTATGTTGACCTAATTAGG 114752

QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPlysIleuAlaValAlaIleLys 40
|||||
Db 114751 AAGCTGTCAATTATTGATCTCAAGAGGATGGAAGAGTTAGCTGTAGCTATTAAA 114692

QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArg 54
|||||
Db 114691 AAACCATCTGTGATGATAGTATCAATCAGTTTCACATAAGG 114650
```

```
RESULT 15
AX321132 501 bp DNA linear PAT 15-DEC-2001
LOCUS Sequence 149 from Patent WO0177168.
AX321132
VERSION AX321132.1 GI:17904576
KEYWORDS
```

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Lodes, M.J., Wang, T., Mohamath, R. and Indirias, C.Y. Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

JOURNAL Patent: WO 0177168-A 149 18-OCT-2001; CORIXA CORPORATION (US) Location/Qualifiers

```
source 1..501
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 166 a 102 c 97 g 134 t 2 others
ORIGIN
```

```
Alignment Scores:
Pred. No.: 2.97e-29 Length: 501
Score: 277.00 Matches: 54
Percent Similarity: 98.18% Conservative: 0
Best Local Similarity: 98.18% Mismatches: 1
Query Match: 85.76% Indels: 0
DB: 6 Gaps: 0
```

US-10-001-254-26 (1-59) x AX321132 (1-501)

```
QY 1 MetAsnlySProlleThrProSerThrYrValArgCysLeuAsnValGlyLeuIleArg 20
|||||
Db 21 ATGACAAACCCATACACATCATATCTGCGCTTATGTTGACCTAATTAGG 80

QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPlysIleuAlaValAlaIleLys 40
|||||
Db 81 AAGCTGTCAATTATTGATCTCAAGAGGATGGAAGAGTTAGCTGTAGCTATTAAA 140

QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArg 54
|||||
Db 141 AAACCATCTGTGATGATAGTATCAATCAGTTTCACATAAGG 183
```

Search completed: January 18, 2004, 04:27:07
Job time : 1025.41 secs

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 18, 2004, 01:11:43 / Search time 576.135 Seconds
(without alignments)
2488.940 Million cell updates/sec

Title: US-10-001-254-26
Perfect score: 323
Sequence: 1 MNKPITPSTYVRCINVLIR.....KKPSGDRDYNQPHIRCCSQN 59

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 segs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n.model -DEV=xlh
-Q=cg2_1/USPTO.spool/US1000154/runat_16012004_152424_19723/app_query.fasta_1.1109
-DB=EST -QPM=fastcap -SUFFIX=ext -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=prc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US1000154.@CCN_1.1_4382.@runat_16012004_152424_19723 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEODDERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vtl:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_pbg:*
27: em_gss_vtl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	323	100.0	284	9	AA114228 zn75g05.r
2	296	91.6	576	10	BF238344 601904613
3	290	89.8	540	9	AL699213 DXF2P686K
4	290	89.8	719	10	BF66981 602130160
5	290	89.8	811	10	BG616438 602642772
6	290	89.8	859	10	BG164491 602342026
7	284	87.9	313	12	BM151935 TCBAPIE11
8	282	87.3	265	9	AW436511 76774 MAR
9	274	84.8	402	10	BE482619 168463 BA
10	271	83.9	453	10	BB860349 BB860349
11	271	83.9	503	14	CA538859 C0272B01-
12	271	83.9	507	10	BB866698 BB866698
13	271	83.9	524	9	AM106160 umc23h11.y
14	271	83.9	575	4	BX522921 RZPD Mus
15	271	83.9	598	13	BO552228 HA014C09-
16	271	83.9	610	10	BB660378 BB660378
17	271	83.9	637	10	BB613447 BB613447
18	271	83.9	638	14	BY721552 BY721552
19	271	83.9	663	14	BY726858 BY726858
20	271	83.9	676	10	BB613167 BB613167
21	271	83.9	1161	11	AK020397 Mus muscu
22	271	83.9	2481	11	AK028837 Mus muscu
23	271	83.9	2810	11	AK029028 Mus muscu
24	269	83.3	888	10	BF687921 602066996
25	197	61.0	629	9	AL647125 AL647125
26	197	59.4	520	10	BE132064 db41h09.y
27	192	58.8	664	12	BJ035962 BJ035962
28	190	58.8	882	13	B0209111 B0209111
29	186	57.6	600	9	AJ453616 AJ453616
30	186	57.6	670	9	AJ447581 AJ447581
31	159	49.2	541	10	BG691069 BG691069
32	135	41.8	632	14	CA365604 CA365604
33	125	38.7	555	9	AM423082 f68b07.y
34	125	38.7	743	14	CA474136 AGENCOURT
35	111	34.4	318	14	CB940191 lPCGx14-
36	100	31.0	380	12	BI863790 FN07C09.y
37	74	22.9	477	10	AM958112 EST370182
38	73.5	22.8	584	10	BG078473 H3028C02-
39	73.5	22.8	588	12	BM022353 i659c09.y
40	73.5	22.8	884	12	BI854371 BI854371
41	71	22.0	508	12	BJ050938 BJ050938
42	69	21.4	605	12	BJ499048 BJ499048
43	69	21.4	631	12	BM138979 HC ad 15D
44	68.5	21.2	778	14	CB236976 AGENCOURT
45	68	21.1	803	28	BH443382 B0GGL71TR

DNA encoding 284 bp

ALIGNMENTS

RESULT 1
AA114228
LOCUS
DEFINITION zn75g05.r1 Striatogene NT2 neuronal precursor 937230 Homo sapiens
ACCESSION AA114228
VERSION AA114228
KEYWORDS CDNA clone IMAGE:564056 5', mRNA sequence.
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 284)

AUTHORS

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Maris, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellendy, K., Soares, M.B., Tan, F., Thierly-Meg, V., Trevisan, E., Underwood, K., Wohlman, P., Waterston, R., Wilson, R. and Warr, M. Generation and analysis of 280,000 human expressed sequence tags Genome Res. 6 (9), 807-828 (1996)

JOURNAL MEDLINE PUBMED

97044478
889549

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LILN; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from AmerSham.

FEATURES

source

Location/Qualifiers
1..284
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:4596579"
/db_xref="taxon:9606"
/clone="IMAGE:564056"
/issue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/note="Organ: brain; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Uninduced, exponentially growing neuroepithelial cells (Ntera-2/cl.D1). Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGCGCAG 3' ~3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"

BASE COUNT 93 a 63 c 56 g 71 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 2,76e-36 Length: 284
Score: 323.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-001-254-26 (1-59) x AA114228 (1-284)

QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
DB 61 ATGAACAAACCCTAATACACATATGTCGCTGCTCAATGTGAGCAATATTAGG 120

QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPylsLeuAlaValAlaIleLys 40
DB 121 AAGCTGTCAAGATTTTATGATCTCAAGAGATGAGAAAGTTAGCTGATTAATAA 180

QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgCysCysSerGlnAsn 59
DB 181 AAACCATCTGGTATGATATCAATCAATGTTTCACATAGATGCTGTTCCAAAC 237

RESULT 2

BF238344

LOCUS BF238344 576 bp mRNA linear EST 14-NOV-2000
DEFINITION BF238344 Homo sapiens cDNA clone IMAGE:4132682 5',
mRNA sequence.

ACCESSION BF238344
VERSION BF238344.1 GI:11152264
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 576)
AUTHORS NIH-MGC <http://mcc.nhl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)
DNA Sequencing-by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILN at: <http://image.llnl.gov>
Plate: LILN1032 row: 1 column: 03
High quality sequence start: 6
High quality sequence stop: 536.
Location/Qualifiers

FEATURES

source

1..576
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4132682"
/issue_type="from chronic myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 54"
/note="Organ: bone marrow; Vector: pDNR-III (Clontech); Site 1: SfiI (ggccatcggcc); Site 2: SfiI (ggccatcggcc); Double stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATCTAGAGCGCGCGCGGACACTG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 191 a 108 c 132 g 145 t
ORIGIN

Alignment Scores:

Pred. No.: 5,43e-32 Length: 576
Score: 296.00 Matches: 54
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.18% Mismatches: 0
Query Match: 91.64% Indels: 0
DB: Gaps: 0

US-10-001-254-26 (1-59) x BF238344 (1-576)

QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
DB 68 ATGAACAAACCCTAATACACATATGTCGCTGCTCAATGTGAGCAATATTAGG 127

QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPylsLeuAlaValAlaIleLys 40
DB 128 AAGCTGTCAAGATTTTATGATCTCAAGAGATGAGAAAGTTAGCTGATTAATAA 187

QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgCys 55
DB 188 AAACCATCTGGTATGATATCAATCAATGTTTCACATAGATGCTGTTCCAAAC 232

RESULT 3

AL699213

LOCUS AL699213 540 bp mRNA linear EST 21-MAR-2002
DEFINITION DKF2p686K18112.r1 686 (synonym: hicc3) Homo sapiens cDNA clone
mRNA sequence.

ACCESSION AL699213
VERSION AL699213.1 GI:19619753
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 540)
 AUTHORS Poustka, A., Wellenreuther, R., Mewes, H.W., Well, B. and Wiemann, S.).
 TITLE EST (Poustka, A., Wellenreuther, R., Mewes, H.W., Well, B. and Wiemann, S.)
 JOURNAL COMMENT Unpublished
 Contact: Poustka A.J.
 Department Leirach
 Max-Planck-Institute for Molecular Genetics
 Innesstrasse 73, 14195 Berlin, Germany
 Tel: +49-30-84131623
 Fax: +49-30-84131128
 Email: poustka@mpg-berlin-dahlem.mpg.de
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by DKFZ (German Cancer Research Center,
 Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.
 No si sequence available.
 This clone (DKFZp686K18112) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
 source
 1. 540
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp686K18112"
 /class_type="human skeletal muscle"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="686 (synonym: hlcc3)"
 /note="Vector: pTriplex2; Site_1: SfiI; Site_2: SfiI; cDNA-collection"

BASE COUNT 150 a 104 c 111 g 174 t 1 others

ORIGIN

Alignment Scores:
 Pred. No.: 3,59e-31 Length: 540
 Score: 290.00 Matches: 54
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 89.78% Indels: 0
 DB: 9 Gaps: 0

US-10-001-254-26 (1-59) x AL699213 (1-540)

QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
 Db 61 ATGAACAACCCATACACCATCAATATGTGGCTGCCTCATGTGGACTAATTAGG 120

QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTTrpLysLysLeuAlaValAlaIleLys 40
 Db 121 AAGCTGTCAAGATTTTATGATCTCTCAAGAGATGAGAGAAAGTATGCTGATATTAA 180

QY 41 LysProSerGlyAspAspArgTyTyrAsnGlnPheHisIleArg 54
 Db 181 AAACCATCTGGTGTATGATACATCAATCAAGTTTCACTAATAGG 222

RESULT 4
 BF696981 719 bp mRNA linear EST 22-DEC-2000
 LOCUS 60210160F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4287014 5',
 DEFINITION mRNA sequence.
 ACCESSION BF696981 GI:11982389
 VERSION BF696981.1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 719)
 NIH-MGC http://mgs.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LNCM124 row: 9 column: 15
 High quality sequence stop: 632.

FEATURES
 source
 1. 719
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4287014"
 /issue_type="primitive neuroectoderm"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_56"
 /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgctcgcc); Site_2: SfiI (ggccatagggc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTTGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGAGCGGCGCATG-3' (30)BN-3' (where B = A, C, G or N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 216 a 145 c 163 g 195 t

ORIGIN

Alignment Scores:
 Pred. No.: 5,34e-31 Length: 719
 Score: 290.00 Matches: 54
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 89.78% Indels: 0
 DB: 10 Gaps: 0

US-10-001-254-26 (1-59) x BF696981 (1-719)

QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
 Db 68 ATGAACAACCCATACACCATCAATATGTGGCTGCCTCATGTGGACTAATTAGG 127

QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTTrpLysLysLeuAlaValAlaIleLys 40
 Db 128 AAGCTGTCAAGATTTTATGATCTCTCAAGAGATGAGAGAAAGTATGCTGATATTAA 187

QY 41 LysProSerGlyAspAspArgTyTyrAsnGlnPheHisIleArg 54
 Db 188 AAACCATCTGGTGTATGATACATCAATCAAGTTTCACTAATAGG 229

RESULT 5
 BG616438 811 bp mRNA linear EST 18-APR-2001
 LOCUS 602642772F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4773760 5',
 DEFINITION mRNA sequence.
 ACCESSION BG616438
 VERSION BG616438.1 GI:13667809
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 811)
 NIH-MGC http://mgs.nci.nih.gov/.

NIH-MGC http://mgs.nci.nih.gov/.

1102 Bates, MC3-3320 Houston, TX 77030, USA
 Tel: 832-824-4536
 Fax: 832-825-4038
 Email: clones@txccc.org
 Seq primer: M13 primer
 Location/Qualifiers

FEATURES

source

```
1. .313
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TCBAP1145"
/sex="male"
/tissue_type="leukophoresis"
/cell_type="pre-B cell"
/dev_stage="pediatric 2 years"
/lab_host="MDH10B"
/clone_lib="pediatric pre-B cell acute lymphoblastic
leukemia Baylor-HSC project-TCBA"
/note="Vector: lambda pSB; Site 1: BamHI; Site 2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGGACGCGGCGGAGGAGG(7)VN
3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second strand
was primed with a BamHI-dC primer
[5'AGAGGCTCGATCGGCGCGCATATATATAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Garinici P, Westover A, Nishiyama Y, Ohsumi T,
Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)"
```

BASE COUNT 82 a 62 c 81 g 87 t 1 others

ORIGIN

Alignment Scores:
 Pred. No.: 1,22e-30 Length: 313
 Score: 284.00 Matches: 53
 Percent Similarity: 98.15% Conservative: 0
 Best Local Similarity: 98.15% Mismatches: 1
 Query Match: 87.93% Indels: 0
 DB: 12 Gaps: 0

US-10-001-254-26 (1-59) x BM151935 (1-313)

QY 1 MetAsnLysProIleThrProSerThyTyrValArgCysLeuAsnValGlyLeuIleArg 20
 |||||
 DB 53 ATGAACAACCCATAATACCTTCAACATATGTGGCTGCTCAATGTTGACCTAATTAGG 112
 |||||
 QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPheLysLeuAlaValAlaIleLys 40
 |||||
 DB 113 AAGCTGTCAATTTATTATGATCCCAAGAGATGGAAGATTAGCTGTAGCTATTAA 172
 |||||
 QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArg 54
 |||||
 DB 173 AAACCATCTGCTGATGATGATACATCAATCAGTTTCACATAAGG 214
 |||||

RESULT 8 265 bp mRNA linear EST 09-JUL-2000
 AM436511
 LOCUS 76774 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
 DEFINITION AM436511
 ACCESSION AM436511
 VERSION AM436511.1 GI:6971817
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa

REFERENCE
 AUTHORS Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
 Vallet,J., Wise,T., Rohrer,G.A., Perlee,G., Sultana,R., Quackenbush
 J., and Keele,J.W.

TITLE Porcine gene discovery by normalized cDNA-library sequencing and
 EST cluster assembly
 JOURNAL Mamm. Genome 13 (8), 475-478 (2002)
 MEDLINE 22213789
 PUBMED 12226715
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smtlh@mail.marc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 20
 and -mismatch 12 options.
 PCR primers
 FORWARD: AGCAACACGCTATGACCAT
 BACKWARD: GTTTCGCCGACGACG
 Plate: 31 row: H column: 17
 Seq primer: ATTAGGTGACCTATAG.
 Location/Qualifiers

FEATURES

source

```
1. .265
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="MDH10B"
/clone_lib="MARC 2P1G"
/note="Vector: pCMV SPORTE; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from testis, ovary,  
endometrium, hypothalamus, pituitary, and placenta."
```

BASE COUNT 90 a 53 c 58 g 64 t

ORIGIN

Alignment Scores:
 Pred. No.: 1,87e-30 Length: 265
 Score: 282.00 Matches: 53
 Percent Similarity: 98.15% Conservative: 0
 Best Local Similarity: 98.15% Mismatches: 1
 Query Match: 87.31% Indels: 0
 DB: 9 Gaps: 0

US-10-001-254-26 (1-59) x AM436511 (1-265)

QY 1 MetAsnLysProIleThrProSerThyTyrValArgCysLeuAsnValGlyLeuIleArg 20
 |||||
 DB 32 ATGAACAACCCATAATACCTTCAACATATGTGGCTGCTCAATGTTGACCTAATTAGG 91
 |||||
 QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrPheLysLeuAlaValAlaIleLys 40
 |||||
 DB 92 AAGTGTCAATTTATTATGATCCCTCAAGAGATGGAAGATTAGCTGTAGCTATTAA 151
 |||||
 QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArg 54
 |||||
 DB 152 AAACCATCTGCTGATGATGATACATCAACAGTTTCATATAAGG 193
 |||||

RESULT 9 402 bp mRNA linear EST 27-MAR-2003
 BE482619
 LOCUS 168463 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION BE482619
 ACCESSION BE482619
 VERSION BE482619.1 GI:9602152
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus

REFERENCE
 AUTHORS Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 Sonstegard,T., Cepuco,A.V., White,J., Van Tassel,C.P., Connor,E.E.,
 Cho,J., Sultana,R., Shade,L., Wray,J.E., Wells,K.D. and
 Quackenbush, J.
 Analysis of bovine mammary gland EST and functional annotation of
 the Bos taurus gene index

JOURNAL Mamm. Genome 13 (7), 373-379 (2002)
 MEDLINE 22135956
 PUBMED 12140686
 COMMENT Contact: Sonstegard TS
 USDA, ARS, Beltsville Agricultural Research Center
 Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
 Tel: 301 504 8416
 Fax: 301 504 8414
 Email: tad@psl.barc.usda.gov
 Single pass sequencing. Bases called and alt. trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCGACGACGACG
 Plate: 11 row: F column: 7
 Seq primer: ATTTAGGTGACCTATAG.
 Location/Qualifiers
 1. 402
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /feature_type="pooled"
 /lab_host="DH10B"
 /clone_lib="BARC SBOV"
 /note="Vector: PCMV SPOR6; Site 1: NotI; Site 2: SalI;
 library made from pooled mRNA isolated from mammary
 tissues at eight physiological, developmental, and disease
 states."
 BASE COUNT 114 a 95 c 88 g 105 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,676-29 Length: 402
 Score: 274.00 Matches: 51
 Percent Similarity: 98.15% Conservative: 2
 Best Local Similarity: 94.44% Mismatches: 1
 Query Match: 84.83% Indels: 0
 DB: 10 Gaps: 0
 US-10-001-254-26 (1-59) x BE482619 (1-402)
 QY 1 MetAsnLysProIleThrProSerThyTyValArgCysLeuAsnValGlyLeuIleArg 20
 DB 64 ATGAACAAAGCCGTTGACATGACATGACATGACGCGCTGCTGGACCAATTTAG 123
 QY 21 LysLeuSerAspHeiLeaSpProGlnGlnGlyTTrpLysLysLeuAlaValAlaIleLys 40
 DB 124 AAGCTGTCCGATTTTATTTGATCCTCAAGAGATGAGAGATTAGCAGTCGCTATTAA 183
 QY 41 LysProSerGlyAspAspArgTyTyrAsnGlnPheHisIleArg 54
 DB 184 AAACCATCTGGTATGATGATGATCAATCAATGTTTCACATTAAG 225
 RESULT 10 BB860349 453 bp mRNA linear EST 26-NOV-2001
 BB860349 RIKEN full-length enriched, pooled cell lines Mus musculus
 LOCUS BB860349
 DEFINITION CDNA clone G430013H14 5', mRNA sequence.
 ACCESSION BB860349
 VERSION BB860349.1 GI:17101803
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 453)
 Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
 Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishii,
 Y., Ito,M., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Matsuyama,T.,
 Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
 Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
 Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa

TITLE
 JOURNAL
 COMMENT Unpublished
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Shuhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
 M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Mochizuki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
 S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
 Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multichipillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Kono,H., Fukushima,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
 Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.
 e mouse tissues.
 Location/Qualifiers
 1. 453
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone_lib="G430013H14"
 /clone_lib="RIKEN full-length enriched, pooled cell lines"
 /note="pooled cell lines; (cell_line=CR1-1751 WEHI 164),
 (cell_line=CR1-2116 JC), (cell_line=RCB-0035 WEHI-3),
 (cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHTA),
 (cell_line=RCB-0559 K-1, FI), (cell_line=RCB-1283 B16
 melanoma), (cell_type=B cells, cell_line=CR1-1702 WEHI 231
), (cell_type=B-lycid cells, cell_line=CR1-2065 M12C-1),
 (cell_type=Nullipotent stem cell, cell_line=CR1-2070 NE),
 (tissue_type=bladder, cell_line=RCB-0544 MBT-2),
 (tissue_type=bone marrow, cell_type=stroma cell,
 cell_line=CR1-2028 SR-4987), (tissue_type=colon,
 cell_line=RCB-0549 C1e-H3), (tissue_type=kidney,
 cell_line=CC1-142 RAG), (tissue_type=submandibular gland,
 cell_line=CC1-1734 SCA-9 Clone 15), (strain=BALB/C,
 cell_type=B cells, cell_line=CR1-1669 BCL1 Clone 13, 20-3B3
), (strain=C3H, tissue_type=brain, cell_line=CR1-1443
 BC3H1)"
 BASE COUNT 110 a 121 c 133 g 89 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,486-28 Length: 453
 Score: 271.00 Matches: 49
 Percent Similarity: 98.15% Conservative: 4
 Best Local Similarity: 90.74% Mismatches: 1
 Query Match: 83.90% Indels: 0
 DB: 10 Gaps: 0
 US-10-001-254-26 (1-59) x BB860349 (1-453)
 QY 1 MetAsnLysProIleThrProSerThyTyValArgCysLeuAsnValGlyLeuIleArg 20
 DB 165 ATGAACAAAGCCGTTGACATGACATGACATGACGCGCTGCTGGACCAATTTAG 224

QY 21 LysleuSerAppheileaspProGlnGluGlyTrpLysLysLeuAlaValAlaIleLys 40
 Db 225 AAGCTGTCGATTATTATGATCCTCAGAGAGGGTGGAGAAATTGACGTAGCTATCAAA 284
 QY 41 LysProSerGlyAspAspArgTYAsnGlnPheHisIleArg 54
 Db 285 AAGCGTCCGGCGACGACAGATACATCAGTTCCATTATTAAG 326
 RESULT 11
 CA538859
 LOCUS
 DEFINITION
 CO272801-5N NIA Mouse 7.5-dpc Whole Embryo cDNA Library (Long) Mus
 musculus cDNA clone NIA:CO272801 IMAGE:30017484 5', mRNA sequence.
 CA538859
 VERSION
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 503)
 Piao, Y., Kazguli, G.D., Dudekula, D.B., Qian, Y., Tanaka, T., Lim, M.K.,
 Luo, A. and Ko, M.S.H.
 Systematic Analyses of NIA Mouse 7.5-dpc Whole Embryo cDNA Library
 (Long)
 Unpublished
 Other ESTs: CO272801-3
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6620, USA
 Email: cdna@gsun.grc.nia.nih.gov
 Plate: CO272 row: B column: 01
 Seq primer: M13 Reverse
 High quality sequence stop: 503
 POLYA=No.
 FEATURES
 SOURCE
 Location/Qualifiers
 1..503
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /db_xref="EAST:CO272801-5N"
 /clone="NIA:CO272801 IMAGE:30017484"
 /tissue_type="whole embryo including extraembryonic
 tissues at 7.5-days postcoitum"
 /dev_stage="7.5-days postcoitum"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse 7.5-dpc Whole Embryo cDNA Library
 (Long)"
 /note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is
 a long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
 extracted from a pool of four embryos at 7.5-days
 postcoitum. Double-stranded cDNAs were synthesized with an
 Oligo(dT) primer [Invitrogen:
 5'-GACTAGTTCTAGATCGAGCGCCGCCCTTTT-3'] from
 7 ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to Lome-linker Lr-Sal4, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Tag polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pSPORT1 plasmid vector.
 The DH10B E. coli host was transformed with the ligation
 mixture by the standard chemical method. The average
 insert size is about 2.2 kb. The library was constructed

BASE COUNT 119 a 146 c 138 g 100 t
 ALIGNMENT SCORES:
 Pred. No.: 1,71e-28 Length: 503
 Score: 271.00 Matches: 49
 Percent Similarity: 98.15% Conservative: 1
 Best Local Similarity: 90.74% Mismatches: 4
 Query Match: 83.90% Indels: 0
 DB: 14 Gaps: 0
 US-10-001-254-26 (1-59) x CA538859 (1-503)
 QY 1 MetasnllyProlleThrrProSerThrTYValAlrGcySleuAenValGlyLeuIleArg 20
 Db 168 ATGACCAAGCGCGTGGACACATCAGCATACATCAGCAACTTAATGCGGATCCTTAAG 227
 QY 21 LysleuSerAppheileaspProGlnGluGlyTrpLysLysLeuAlaValAlaIleLys 40
 Db 228 AAGCTGTCGATTATTATGATCCTCAGAGAGGGTGGAGAAATTGACGTAGCTATCAAA 287
 QY 41 LysProSerGlyAspAspArgTYAsnGlnPheHisIleArg 54
 Db 288 AAGCGTCCGGCGACGACAGATACATCAGTTCCATTATTAAG 329
 RESULT 12
 BB866698
 LOCUS
 DEFINITION
 BB866698 RIKEN full-length enriched, pooled cell lines Mus musculus
 cDNA clone G4D0001K20 5', mRNA sequence.
 BB866698
 VERSION
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 507)
 Akimura, T., Arikawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
 Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imotani, K., Ishii
 Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T.,
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T.,
 Saito, R., Sakai, C., Sakai, K., Sakazume, N., Saeki, D., Sato, K.,
 Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa
 A., Takehashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T.,
 Watanishi, A., Yasunishi, A., Yamatsu, M., and Hayashizaki, Y.
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.
 2001)
 Unpublished
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-2 Sphiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multichannel sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 Y. and Hayashizaki, Y.


```

XX EST, expressed sequence tag.
KM
XX Mus musculus (house mouse)
OS
XX Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC
XX Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
XX
XX [1]
RP 1-575
RA Heil O., Ebert L., Neubert P., Peters M., Radelof U., Schneider D.,
RT Korn B.;
RL Submitted (28-MAY-2003) to the EMBL/GenBank/DBJ databases.
RL RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Im Neuenheimer
RL Feld 580, D-69120 Heidelberg, Germany
XX
CC RZPD; IMAGP998L225504.
CC RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
CC Mouse Unigeneset - RZPD2 (RZPDLIB No.981)
CC http://www.rzpd.de/CloneCards/cgi-bin/showlib.pl.cgi/response?libNo=981
CC Contact: Ina Kollis
CC RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
CC Heubnerweg 6, D-14059 Berlin, Germany
CC Tel: +49 30 32639 101
CC Fax: +49 30 32639 111
CC www.rzpd.de
CC This clone is available royalty-free from RZPD;
CC contact RZPD (clone@rzpd.de) for further information.
CC Seq primer: sugf, Primer sequence: CTTGCTGCTMAAGCTGCG
XX
XX Key Location/Qualifiers
FH
FH source
FH 1. 575
FH /db_xref="taxon:10090"
FH /note="1st strand cDNA was primed with an oligo(dT) primer
FH [ATGCGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was
FH ligated to a Dralli adaptor [GTTGGCGCTACTGC], digested and
FH cloned into distinct Dralli sites of the pME18S-FL3 vector
FH (5' site CACTGTGTC, 3' site CACCATGTG). XhoI should be used
FH to isolate the cDNA insert. Size selection was performed to
FH exclude fragments <1.5kb. Library constructed by Dr. Sumio
FH Sugano (University of Tokyo Institute of Medical Science).
FH Custom primers for sequencing: 5' end primer
FH CTTGCTGCTMAAGCTGCG and 3' end primer
FH CGACCTCGACCTCGACGAC. REFERENCES: Suzuki, Y., Yoshitomo,
FH K., Maruyama, K., Suyama, A., and Sugano, S. Construction
FH and characterization of a full length-enriched and a 5' end
FH enriched cDNA library. Gene 200, 145-156, 1997. Saeki, Z.,
FH Suzuki, Y., Watanabe, M., Imai, J., Shibui, A., Yoshida,
FH K., Hata, H., Yamaguchi, R., Tateyama, S., and Sugano, S.
FH Construction of mouse full length-enriched cDNA libraries
FH by oligo-capping. DNA Research, submitted."
FH /organism="Mus musculus"
FH /clone="IMAGP998L225504"
FH /clone_id="Sugano mouse embryo mewa"
FH /dev_stage="embryo, 14 dpc"
FH /lab_host="DH10B"
XX
XX Sequence 575 BP; 132 A; 172 C; 155 G; 114 T; 2 other;
SO

```

Alignment Scores:

Pred. No.:	2.06e-28	Length:	575
Score:	271.00	Matches:	49
Percent Similarity:	98.15%	Conservative:	4
Best local Similarity:	90.74%	Mismatches:	1
Query Match:	83.90%	Indels:	0
DB:	4	Gaps:	0

US-10-001-254-26 (1-59) x BXS22921 (1-575)

```

Oy 1 MetanlyspProlleThThrValArgCysleuAnValGlyleuIleArg 20
Db 216 ATGAACAAGCCGTTGACACCAATGCACATTAATGATGAGGATCCTTAG 275

```

Qy	21	LysleusearpprellieaspProGlnGly1TrpLysLysleualaValAlaIleLys	40
Db	276	AAAGTCGCGATTTTATTGATCTCTCAAGAGGGTGGAGGAATTAGCATGACTATCAAA	335
Qy	41	LysProSerGlyAaPaAParGlyTyrAsnGlnPheHisIleArg	54
Db	336	AAAGCGTCGGCGACGACAGATACATCACTTCATATTAAG	377
RESULT 15			
LOCUS	BQ552228		
DEFINITION	H4014C09-5 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone		
ACCESSION	H4014C09 5', mRNA sequence.		
VERSION	BQ552228.1	GI:21453114	
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.		
REFERENCE	1 (bases 1 to 598)		
AUTHORS	Vanburen,V., Piao,Y., Dudekula,D.B., Qian,Y., Carter,M.G., Martin P.R., Stagg,C.A., Baasey,U., Alba,K., Hamatani,T., Kargul,G.J., Luo,A.G., Kelso,J., Hide,W. and Ko,M.S.H.		
TITLE	Assembly, verification, and initial annotation of NIA 7.4K mouse cDNA clone set		
JOURNAL	Genome Res. 12 (12), 1999-2003 (2002)		
MEDLINE	22354164		
PUBMED	12466305		
COMMENT	Contact: Yong Qian Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA Email: cchna@genun.grc.nia.nih.gov This clone set has been freely distributed to the community. Please visit http://lgun.grc.nia.nih.gov/cDNA/NIA_7_4K.html for details. Plate: H4014 row: C column: 09 Seq primer: -21M13 Reverse High quality sequence stop: 598 POLYA=NO.		
FEATURES			
source	Location/Qualifiers 1..598 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6" /db_xref="taxon:10090" /db_xref="taxon:10090" /clone="H4014C09" /sex="mixed" /dev_stage="mixed" /lab_host="DH10B" /clone_1lb="NIA Mouse 7.4K cDNA Clone Set" /note="Vector: pSPORT1, Site 1: SalI; Site 2: NotI; This clone is among a rearayted set of 7,407 clones from more than 20 cDNA libraries."		
BASE COUNT	150 a 174 c 163 g 111 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,18e-28	Length:	598
Score:	271.00	Matches:	49
Percent Similarity:	98.15%	Conservative:	4
Best Local Similarity:	90.74%	Mismatches:	1
Query Match:	83.90%	Indels:	0
DB:	13	Gaps:	0
US-10-001-254-26 (1-59) x BQ552228 (1-598)			
Qy	1	MetAsnLysProIleThrProSerThrTyrValArgCysLeuAsnValGlyLeuIleArg	20
Db	152	ATGAAACAAGCGGTGTGACACATGACATACATGCGAACCTTAATGATGGGATCTTTAGG	211

```
Qy      21  LysLeuSerAspPheIleAspProGlnGluGlyTyrPlySLysLeuAlaValAlaIleLys 40
Db      212  AAGCTGTCGGATTCTTATGATCCTCAAGAGGGTGGAGAAATTAGCAGTAGCTATCAAA 271
Qy      41  LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArg 54
Db      272  AAGCGTCCGGCGACAGATACATCAGTTCATATAAG 313
```

Search completed: January 18, 2004, 06:06:52
Job time : 578.135 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 18, 2004, 00:40:22 ; Search time 70.7617 Seconds
(without alignments)
2250.748 Million cell updates/sec

Title: US-10-001-254-26
Perfect score: 323
Sequence: 1 MNKRITPSTYVRCNLNGLIR.....KKPSGDDRRYNGFHRCSSQN 59

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Delop 6.0 , Delext 7.0

Searched: 2552756 segs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgm2_1/USP10.spool/US10001254/runat_16012004_152423_19704/apd_query.fasta_1.1109
-DB=N-Geneseq_19Jun03 -QPM=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOFC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=b1osum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=PLC -NOEM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US10001254.@CGN_1.1.0.@runat_16012004_152423_19704 -NCPU=6 -ICPU=3
-NO MMAP -LARGEDBURY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N-Geneseq_19Jun03: *
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	323	100.0	211	24	AAD40084	Human IRAK4 short
2	323	100.0	415	23	AA576802	DNA encoding novel
3	323	100.0	2213	22	AAH13798	Human cDNA sequenc
4	290	89.8	833	21	AA09319	Human cancer assoc
5	290	89.8	1383	22	AAD10197	Human interleukin-
6	290	89.8	1383	24	AAD40079	Human IRAK4 gene #
7	290	89.8	1668	23	AA576805	DNA encoding novel
8	290	89.8	2817	24	AAAD0085	Human IRAK4 gene #
9	277	85.8	501	24	AA561608	Lung small cell ca
10	271	83.9	1542	22	AAD10198	Mouse interleukin-
11	247	76.5	294	24	AAD40074	Human IRAK4 DP (de
12	68	21.1	5170	21	AA016683	Human secreted pro
13	68	21.1	349980	24	AB081849	Bifidobacterium 10
14	66	20.4	972	23	ABL27971	Drosophila melanog
15	66	20.4	3610	23	ABL27970	Drosophila melanog
16	65.5	20.3	1257	21	AA033451	Arabidopsis thalia
17	65.5	20.3	4151	22	ABA06564	Human cDNA SEQ ID
18	65.5	20.3	4151	24	ABV83901	Human polynucleoti
19	64.5	20.0	2893	23	ABL25610	Drosophila melanog
20	64	19.8	410	22	AAK56325	Human immune/haema
21	64	19.8	1301	25	ABQ84283	Full length cDNA 1
22	64	19.8	2388	24	ABA05888	Human aminopeptida
23	64	19.8	2391	25	ABQ84286	Mouse DEPI0 coding
24	64	19.8	2583	24	ABK83324	cDNA encoding huma
25	64	19.8	2767	24	ABQ75945	Human PWM encodin
26	64	19.8	2909	25	ABQ84287	Mouse DEPI0 cDNA s
27	64	19.8	2909	25	ABQ84292	Mouse DEPI0 full l
28	64	19.8	3238	24	ABA05887	Human aminopeptida
29	64	19.8	3550	25	ABQ84285	Human DEPI0 cDNA s
30	63.5	19.7	206	22	AA511008	Cone snail cDNA en
31	63	19.5	555	24	ABQ52190	Oligonucleotide fo
32	63	19.5	555	24	ABQ52191	Oligonucleotide fo
33	63	19.5	2944528	24	ABA03041	Listeria monocytog
34	62.5	19.3	200	22	AA511011	Cone snail cDNA en
35	62	19.2	29417	23	ABL13008	Drosophila melanog
36	61.5	19.0	304	22	ABA66305	Human foetal liver
37	61.5	19.0	304	22	ABA33370	Probe #11836 for g
38	61.5	19.0	304	22	AAK14722	Human brain expres
39	61.5	19.0	304	22	AAK40465	Human bone marrow
40	61.5	19.0	304	22	AA121224	Probe #11157 for g
41	61.5	19.0	304	22	AA146496	Probe #15182 used
42	61.5	19.0	304	22	AA106933	Probe #6924 used t
43	61.5	19.0	304	23	AB540029	Human liver single
44	61.5	19.0	304	24	AB514456	Human genome-deriv
45	61.5	19.0	472	23	AB528667	Human liver single

ALIGNMENTS

RESULT 1
ID AAD40084 standard; DNA; 211 BP.
XX AAD40084;
AC
XX
DT 22-OCT-2002 (first entry)
XX
DE Human IRAK4 short gene.
XX
XX
KW Human; death domain; DD; death effector domain; DED; Chlamydia infection;
NM-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
inflammation; allergy; autoimmunity; allograft rejection; cell division;
immune-based pathology; fibrosis; arthritis; graft versus host disease;
immunosuppressive; gene therapy; antisense therapy; gene; ds.
OS Homo sapiens.

```

XX Key Location/Qualifiers
FH CDS 1..180
FT /tag= a
FT /product= "Human IRAK4 short protein"
XX
XX MO200240680-A2.
XX
XX 23-MAY-2002.
XX
XX 15-NOV-2001; 2001WO-US44844.
XX
XX 17-NOV-2000; 2000US-0715893.
XX
XX 29-JUN-2001; 2001US-301889P.
XX
XX (BURN-) BURRHAM INST.
XX
XX Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roch W,
XX Stenner-Jewen F,
XX
XX WPI; 2002-500222/53.
XX
XX P-PSDB; AAE24864.
XX
XX
XX New polypeptide comprising a death domain or death effector domain,
XX useful for discovery of drugs that suppress infection, inflammation,
XX allergy, sepsis, autoimmunity, allograft rejection and other diseases
XX
XX
XX Claim 19; Page 195; 209pp; English.
XX
XX The invention relates to an isolated polypeptide comprising a death
XX domain (DD), death effector domain (DED) or NB-ARC domain. The invention
XX is useful for identifying a binding agent, preferably a protein or a drug
XX that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
XX domain from DAB3, IRAK4, CTDD (Chlamydia trachomatis DD protein), DED4 or
XX NIDD (NGFR-interacting Death Domain), with a candidate binding agent and
XX detecting the association of the domain and the candidate binding agent,
XX by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
XX chemical crosslinking, nuclear magnetic resonance (NMR), mass
XX spectroscopy (MS) and FPA. The invention is useful for modulating the
XX level of a cell process such as cell proliferation, cell adhesion, cell
XX stress responses, responses to microbial infection and B cell
XX immunoglobulin class switching, in particular apoptosis within a cell.
XX Antibody specifically reacting with CTDD DD of C. trachomatis, C.
XX muridarum, C. pneumoniae, and C. peitraci or a nucleic acid encoding the
XX CTDD DD protein is useful for detecting a Chlamydia infection. The
XX invention is useful for modulating the activity of oncogenic proteins,
XX for treating a pathology caused by the oncogenic proteins and for
XX treating bacterial infections by modulating the activity of bacterial
XX proteins. The protein and antibody specific for it are useful for
XX discovery of drugs that suppress infection, inflammation, allergy,
XX sepsis, autoimmunity, allograft rejection and other diseases. The protein
XX is useful for treating immune-based pathologies, pathologies associated
XX with cell division, inflammatory diseases such as sepsis, fibrosis,
XX arthritis, graft versus host disease. The invention is used in antisense
XX therapy and gene therapy. The present sequence is human IRAK4 short gene.
XX
XX
XX Sequence 211 BP; 77 A; 41 C; 36 G; 57 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1.63e-40 Length: 211
XX Score: 323.00 Matches: 59
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 24 Gaps: 0
XX
XX US-10-001-254-26 (1-59) x AAD40084 (1-211)
XX
XX 1 MetAsnLysPProIleThrProSerThrTyValArgCysLeuAlaValGlyLeuIleArg 20
XX 1 ATGAACAAACCCATTAACACCATCATATGTCGCTGCTCAATGTTGACATTAATTAG 60

```

```

OY 21 LysLeuSerAspPheIleAspProGlnGluGlyTyrTlPylsLysLeuAlaValIleLys 40
DB 61 AAGCTCTCGAGTTTATTGATCTCTCAAGAGAGATGGAAGAACTGCTTACTATTAA 120
OY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgCysSerGlnAsn 59
DB 121 AAACCATCTGTCGATGATGATACATCATGTTTCACATAGATGCTGTTCCAAAAC 177
XX
XX RESULT 2
XX AAS76802
XX ID AAS76802 standard; cDNA; 415 BP.
XX
XX AAS76802;
XX
XX 13-FEB-2002 (first entry)
XX
XX DE DNA encoding novel human diagnostic protein #12606.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX OS Homo sapiens.
XX
XX PN MO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US08631.
XX
XX PR 31-MAR-2000; 2000US-0540217.
XX
XX PR 23-AUG-2000; 2000US-0649167.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Drmanac RT, Liu C, Tang YT;
XX
XX DR WPI; 2001-639362/73.
XX
XX DR P-PSDB; ABG12615.
XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX
XX PS Claim 1; SEQ ID NO 12606; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Sequence 415 BP; 144 A; 83 C; 78 G; 110 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 4.04e-40 Length: 415
XX Score: 323.00 Matches: 59

```


XX AA09310-20 are novel genes isolated by SEREX screening from a renal
CC cancer cell line 1973/10.4. The genes encode cancer associated antigen
CC precursors. These gene products are useful in methods for preventing,
CC diagnosing and/or treating disorders, especially cancer. Associated with
CC abnormal expression of human cancer associated antigens. The method
CC comprises contacting a sample from a subject with an agent that
CC specifically binds to the nucleic acid molecule or expression product
CC (or fragment) complexed with a human leukocyte antigen (HLA) molecule
CC and determining the interaction between the agent and the nucleic acid
CC molecule or the expression product as a determination of the disorder.
XX
SQ Sequence 833 BP; 273 A; 155 C; 179 G; 226 T; 0 other;
Alignment Scores:
Pred. No.: 1,23e-34 Length: 833
Score: 290.00 Matches: 54
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.78% Indels: 0
DB: 21 Gaps: 0
US-10-001-254-26 (1-59) x AA09319 (1-833)
QY 1 MetAsmlySProlleThrProSerThrTyValArgCysLeuAsnValGlyLeuileArg 20
DB 50 ATGAACAAACCCATACACCATCATGTCGCTGCTCATGTGACTATTAAG 109
QY 21 LysLeuSerAPhelleaSPProGInGluGlyTTPlysLysLeuAlaValAlaileLys 40
DB 110 AAGCTGTGAGATTATTATGATCTCAAGAGAGATGGAAGATGAGCTGATTAATAA 169
QY 41 LysProSerGlyAspAspArgTyTrAsnGlnPheHisileArg 54
DB 170 AAACCATCTGTGATGATGATACATCAATCACTTTCATTAAG 211
RESULT 5
AADD10197 standard; cDNA; 1383 BP.
XX
AC AADD10197;
XX
DT 24-SEP-2001 (first entry)
XX
DE Human interleukin-1 (IL-1) receptor-associated kinase (IRAK)-4 cDNA.
XX
KW Human; interleukin-1 receptor associated kinase-4; IRAK-4; cytosolic;
KW IL; antibacterial; antiinflammatory; ophthalmological; vasotropic; OPD;
KW chronic obstructive pulmonary disease; neuroprotective; chronic cough;
KW adult respiratory disease syndrome; pulmonary fibrosis; asthma; ARDS;
KW interstitial lung disease; allergic rhinitis; transplant rejection;
KW autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;
KW multiple sclerosis; diabetes; cancer; solid tumour; lymphoma; stroke;
KW cardiovascular disease; atherosclerosis; neurodegenerative disease;
KW sepsis; osteoarthritis; rash; osteoporosis; psoriasis; dermatitis;
KW inflammatory bowel disease; gout; Crohn's disease; ulcerative colitis;
KW Behcet's syndrome; ankylosing spondylitis; IL-1 receptor/Toll receptor;
KW sarcoidosis; transgenic animal; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 1.1383
FT CDS /*tag= a
FT /product= "Human IRAK-4"
XX
XX MO200151641-A1.
XX
XX 19-JUL-2001.
XX
XX 12-JAN-2001; 2001MO-US01171.
XX
XX 13-JAN-2000; 2000US-0176395.
XX

XX
PA (TULIA) TULARIK INC.
XX
XX Wesche H, Li S;
PI
XX
XX WPI; 2001-451660/48.
DR
XX P-PSDB; AAE05398.
XX
PT Novel human interleukin-1 receptor associated kinase polypeptide,
PT useful for identifying modulators of the polypeptide for treating gout,
PT asthma, allergic rhinitis, multiple sclerosis and skin cancer
XX
XX Claim 7; Fig 2; 89pp; English.
XX
XX The present sequence is a cDNA encoding human interleukin (IL)-1 receptor
CC associated kinase (IRAK)-4. IRAK associate with activated IL-1, IL-18
CC and other receptors and act to transduce signals originating from the
CC activated receptors, ultimately leading to a variety of downstream
CC effects such as nuclear factor (NF)-kappaB activation. The IRAK-4
CC inhibitors are useful for treating inflammatory diseases such as
CC pulmonary diseases and diseases of the airway (e.g., adult respiratory
CC disease syndrome (ARDS), chronic obstructive pulmonary disease (COPD),
CC pulmonary fibrosis, interstitial lung disease, asthma, chronic cough
CC or allergic rhinitis), transplant rejection, autoimmune diseases (e.g.,
CC rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis
CC or diabetes), cancer (e.g., solid tumours, skin cancer or lymphoma),
CC cardiovascular diseases (e.g., stroke and atherosclerosis), diseases
CC of the central nervous system (e.g., neurodegenerative disease), CD14
CC mediated sepsis, non-CD14 mediated sepsis, osteoarthritis, osteoporosis,
CC psoriasis, diseases of the skin (e.g., rash, contact dermatitis, atopic
CC dermatitis), inflammatory bowel disease (e.g., Crohn's disease and
CC ulcerative colitis), Behcet's syndrome, ankylosing spondylitis, gout,
CC sarcoidosis and ophthalmic diseases and conditions. The inhibitors of
CC IRAK-4 activity or expression are used to inhibit signal transduction
CC resulting from the activation of an interleukin-1 receptor (IL-1R)/Toll
CC receptor in a cell. They also inhibit the activation of a transcription
CC factor that activates NF-kappaB in the cell. IRAK-4 is used to create a
CC nonhuman transgenic animal which is useful for testing the function of
CC IRAK-4 in vivo to generate models for the study of inflammatory
CC disorders and conditions and for the development of potential treatments
CC for IRAK-4-related inflammatory diseases and conditions. IRAK-4 sequences
CC are also used in gene therapy and in antisense therapy.
XX
SQ Sequence 1383 BP; 463 A; 244 C; 283 G; 393 T; 0 other;
Alignment Scores:
Pred. No.: 2.43e-34 Length: 1383
Score: 290.00 Matches: 54
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.78% Indels: 0
DB: 22 Gaps: 0
US-10-001-254-26 (1-59) x AADD10197 (1-1383)
QY 1 MetAsmlySProlleThrProSerThrTyValArgCysLeuAsnValGlyLeuileArg 20
DB 1 ATGAACAAACCCATACACCATCATGTCGCTGCTCATGTGACTATTAAG 60
QY 21 LysLeuSerAPhelleaSPProGInGluGlyTTPlysLysLeuAlaValAlaileLys 40
DB 61 AAGCTGTGAGATTATTATGATCTCAAGAGAGATGGAAGATGAGCTGATTAATAA 120
QY 41 LysProSerGlyAspAspArgTyTrAsnGlnPheHisileArg 54
DB 121 AAACCATCTGTGATGATGATACATCAATCACTTTCATTAAG 162
RESULT 6
AADD40079 standard; DNA; 1383 BP.
XX
AC AADD40079;
XX

DT	22-OCT-2002	(first entry)
XX		
DE	Human IRAK4 gene #1.	
XX		
XX	Human; death domain; DD; death effector domain; DED; Chlamydia infection	
KW	NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis	
KW	inflammation; allergy; autoimmunity; allograft rejection; cell division;	
KW	immune-based pathology; fibrosis; arthritis; graft versus host disease;	
KX	immunosuppressive; gene therapy; antitense therapy; gene; ds.	
XX		
OS	Homo sapiens.	
XX		
XX	Key	Location/Qualifiers
FT	CDS	1..1383
FT		/*tag= a
FT		/product= "Human IRAK4"
PN	MO200240680-A2.	
XX		
PD	23-MAY-2002.	
XX		
PF	15-NOV-2001; 2001WO-US44844.	
XX		
PR	17-NOV-2000; 2000US-0715893.	
PR	29-JUN-2001; 2001US-301889P.	
XX		
PA	(BURN-) BURNHAM INST.	
XX		
PI	Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;	
PI	Stenner-Jiawen F,	
XX		
DR	WPI: 2002-500222/53.	
XX	P-PSDB; AAE24859.	
XX		
PT	New polypeptide comprising a death domain or death effector domain,	
PT	useful for discovery of drugs that suppress infection, inflammation,	
PT	allergy, sepsis, autoimmunity, allograft rejection and other diseases	
PT	-	
XX		
PS	Claim 19; Page 180-182; 209pp; English.	
XX		
CC	The invention relates to an isolated polypeptide comprising a death	
CC	domain (DD), death effector domain (DED) or NB-ARC domain. The invention	
CC	is useful for identifying a binding agent, preferably a protein or a drug	
CC	that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC	
CC	domain from DAP3, IRAK4, CTDP (Chlamydia trachomatis DD protein), DED4 or	
CC	NIDD (NGFR-interacting Death domain), with a candidate binding agent and	
CC	detecting the association of the domain and the candidate binding agent,	
CC	by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or	
CC	chemical crosslinking, nuclear magnetic resonance (NMR), mass	
CC	spectroscopy (MS) and FPA. The invention is useful for modulating the	
CC	level of a cell processes such as cell proliferation, cell adhesion, cell	
CC	stress responses, responses to microbial infection and B cell	
CC	immunoglobulin class switching, in particular apoptosis within a cell.	
CC	Antibody specifically reactive with CTDP DD of C. trachomatis , C.	
CC	muridarum, C. pneumoniae, and C. plicatus or a nucleic acid encoding the	
CC	CTDP DD protein is useful for detecting a Chlamydia infection. The	
CC	invention is useful for modulating the activity of oncogenic proteins,	
CC	for treating a pathology caused by the oncogenic proteins and for	
CC	treating bacterial infections by modulating the activity of bacterial	
CC	proteins. The protein and antibody specific for it are useful for	
CC	discovery of drugs that suppress infection, inflammation, allergy,	
CC	sepsis, autoimmunity, allograft rejection and other diseases. The protein	
CC	is useful for treating immune-based pathologies, pathologies associated	
CC	with cell division, inflammatory diseases such as sepsis, fibrosis,	
CC	arthritis, graft versus host disease. The invention is used in antisense	
CC	therapy and gene therapy. The present sequence is human IRAK4 gene.	
XX		
XX		
SO	Sequence 1383 BP; 463 A; 243 C; 283 G; 394 T; 0 other;	

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	89.78%	Indels:	0
DB:	24	Gaps:	0

US-10-001-254-26 (1-59) x AAD40079 (1-1383)

QY	1	MetAsnLysProIleLeuTherProSerThyTyValAlaGlySLeuAsnValGlyLeuIleAlaArg	20
Db	1	ATGAAACAACCCCTATACACCATCAACATATGTGGCCCTCCTCAATGATTGGACTAATTAGG	60
QY	21	LysLeuSerAspPheIleAspProGlnGluGlyTyrPheLysLeuAlaValAlaIleLys	40
Db	61	AAAGCTGTCAGATTATTATGATCTCTCAAGAGCATGGAAGAGTTAGCTGTAGCTAATTAAA	120
QY	41	LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArg	54
Db	121	AAACCATCTGGTGATGATATGATACATCACTTTCACATTAAG	162

RESULT 7
 AAS76805
 ID AAS76805 standard; cDNA; 1668 BP.
 AC AAS76805;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #12609.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 XX
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Dmanac RT, Liu C, Tang YT;
 XX
 PI WPI; 2001-639362/73.
 XX
 DR P-PSDB; ABG12618.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 1; SEQ ID NO 12609; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1668 BP; 571 A; 289 C; 332 G; 476 T; 0 other;

Alignment Scores:

Pred. No.:	3,12e-34	Length:	1668
Score:	290.00	Matches:	54
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	89.78%	Indels:	0
DB:	23	Gaps:	0

US-10-001-254-26 (1-59) x AAS76805 (1-1668)

QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
DB 1 ATGAACAAACCCTAACACATCACTATGCGCTGCTCAATGTTGACCTAATTAG 60
QY 21 LysLeuSerAspPheIleAspProGingluGlyTyrLysLysLeuAlaValAlaIleLys 40
DB 61 AAGCTGTCAGATTATTATGATCTCTCAAGAGATGAGAGAGTACCTGAGCTATTAA 120
QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArg 54
DB 121 AAACCATCTGTGATGATGATCAATCAATGTTTCACATTAAG 162

RESULT 8

AAD40085
ID AAD40085 standard; DNA; 2817 BP.

AC AAD40085;

DT 22-OCT-2002 (first entry)

DE Human IRAK4 gene #2.

KW Human; death domain; DD; death effector domain; DED; Chlamydia infection;
KW NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
KW inflammation; allergy; autoimmunity; allograft rejection; cell division;
KW immune-based pathology; fibrosis; arthritis; graft versus host disease;
KW immunosuppressive; gene therapy; antisense therapy; gene; ds.

OS Homo sapiens.

FX Location/Qualifiers

FT CDS 50..1432
FT /tag= a
FT /product= "Human IRAK4"

PN WO200240680-A2.

PD 23-MAY-2002.

PF 15-NOV-2001; 2001WO-US44844.

PR 17-NOV-2000; 2000US-0715893.

PR 29-JUN-2001; 2001US-301889P.

PA (BURR-) BURRHAM INST.

PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roch W,
PI Stenner-Jliwen F;

DR WPI: 2002-500222/53.

DR P-PSDB; AAE24865.

PT New polypeptide comprising a death domain or death effector domain,
PT useful for discovery of drugs that suppress infection, inflammation,
PT allergy, sepsis, autoimmunity, allograft rejection and other diseases

PT -
PS Claim 19; Page 194-196; 209pp; English.

CC The invention relates to an isolated polypeptide comprising a death
CC domain (DD), death effector domain (DED) or NB-ARC domain. The invention
CC is useful for identifying a binding agent, preferably a protein or a drug
CC that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
CC domain from DAP3, IRAK4, CTDP (Chlamydia trachomatis DD protein), DED4 or
CC NIDP (NGFR-interacting Death Domain), with a candidate binding agent and
CC detecting the association of the domain and the candidate binding agent,
CC by yeast two hybrid assay, immunoprecipitation, SP, ultraviolet (UV) or
CC chemical crosslinking, nuclear magnetic resonance (NMR), mass
CC spectroscopy (MS) and PDA. The invention is useful for modulating the
CC level of a cell process such as cell proliferation, cell adhesion, cell
CC stress responses, responses to microbial infection and B cell
CC immunoglobulin class switching, in particular apoptosis within a cell.
CC Antibody specifically reactive with CTDP DD of C. trachomatis, C.
CC muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the
CC CTDP DD protein is useful for detecting a Chlamydia infection. The
CC invention is useful for modulating the activity of oncogenic proteins,
CC for treating a pathology caused by the oncogenic proteins and for
CC treating bacterial infections by modulating the activity of bacterial
CC proteins. The protein and antibody specific for it are useful for
CC discovery of drugs that suppress infection, inflammation, allergy,
CC sepsis, autoimmunity, allograft rejection and other diseases. The protein
CC is useful for treating immune-based pathologies, pathologies associated
CC with cell division, inflammatory diseases such as sepsis, fibrosis,
CC arthritis, graft versus host disease. The invention is used in antisense
CC therapy and gene therapy. The present sequence is human IRAK4 gene.

SQ Sequence 2817 BP; 912 A; 547 C; 586 G; 772 T; 0 other;

Alignment Scores:

Pred. No.:	6,32e-34	Length:	2817
Score:	290.00	Matches:	54
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	89.78%	Indels:	0
DB:	24	Gaps:	0

US-10-001-254-26 (1-59) x AAD40085 (1-2817)

QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
DB 50 ATGAACAAACCCTAACACATCACTATGCGCTGCTCAATGTTGACCTAATTAG 109
QY 21 LysLeuSerAspPheIleAspProGingluGlyTyrLysLysLeuAlaValAlaIleLys 40
DB 110 AAGCTGTCAGATTATTATGATCTCTCAAGAGATGAGAGAGTACCTGAGCTATTAA 169
QY 41 LysProSerGlyAspAspArgTyrAsnGlnPheHisIleArg 54
DB 170 AAACCATCTGTGATGATGATCAATCAATGTTTCACATTAAG 211

RESULT 9

AAS61608
ID AAS61608 standard; cDNA; 501 BP.

AC AAS61608;

DT 29-JAN-2002 (first entry)

DE Lung small cell carcinoma antigen, cDNA #149.

KW Human; cytosolic; antitumour; lung small cell cancer antigen;
KW tumour; lung cancer; ss.

OS Homo sapiens.

PN WO200177168-A2.

PT 18-OCT-2001.

XX 11-APR-2001; 2001WO-US11859.
 PF 11-APR-2000; 2000US-196780P.
 XX 21-JUN-2000; 2000US-213361P.
 PR 01-SEP-2000; 2000US-229763P.
 PR 05-SEP-2000; 2000US-230629P.
 PR 14-SEP-2000; 2000US-232565P.
 PR 19-DEC-2000; 2000US-257037P.
 PR 08-JAN-2001; 2001US-260796P.
 XX (CORI-) CORIXA CORP.
 PA
 PI Lodes MJ, Wang T, Mohamath R, Indirias CY;
 DR WPI; 2002-010896/01.
 XX Lung tumour polynucleotide and polypeptides useful in therapy and
 PT diagnosis of cancer especially lung cancer
 XX
 PS Claim 1; Page 174; 295pp; English.
 XX
 CC The invention relates to novel isolated lung small cell cancer antigen
 CC polynucleotides (I) and polypeptides (II) used in a method of detecting
 CC cancer in a patient. The method is optionally performed by
 CC utilising oligonucleotides (III), where the biological sample
 CC from the patient is contacted with (III), detecting the amount of
 CC polynucleotide hybridised to (III) in the sample and comparing the
 CC amount of polynucleotide to a predetermined cut-off value and thereby
 CC determining cancer in a patient. (I), (II) or antigen-presenting cells
 CC expressing (II) is useful for stimulating and/or expanding T cells
 CC specific for a tumour protein. The method comprises contacting T cells
 CC with one of the components under conditions to permit the stimulation
 CC and/or expansion of the cells. A composition comprising (I) is useful for
 CC stimulating an immune response in a patient and for inhibiting the
 CC development of a cancer especially lung cancer in a patient. An
 CC isolated T cell population is useful for removing tumour cells from the
 CC biological sample and for inhibiting the development of cancer in a
 CC patient. AAS61460-AAS61874 represent novel human lung small cell
 CC cancer antigen coding sequences of the invention.
 XX
 SQ Sequence 501 BP; 166 A; 102 C; 97 G; 134 T; 2 other;
 Alignment Scores:
 Pred. No.: 6.18e-33 Length: 501
 Score: 277.00 Matches: 54
 Percent Similarity: 98.18% Conservative: 0
 Best Local Similarity: 98.18% Mismatches: 0
 Query Match: 85.76% Indels: 1
 DB: 24 Gaps: 0
 US-10-001-254-26 (1-59) x AAS61608 (1-501)
 QY 1 MetAsnAnlySProllethnProserThrTyValArgCysLeuAanValGlyLeuIlEArg 20
 Db 21 ATGAACAACCCATTAACCATCAATATGTGGCTCCCTCATTTGGACTAATTAGG 80
 QY 21 LysLeuSerAspPheIlleAspPProGInGluGlyTyrPylsYlsLeuAlaValAlaIlleYs 40
 Db 81 AAGCTGTCAATTTATTTATGATCCCTCAAGAAAGATGGAAGAAGTTAGCTGTAGCTATTAA 140
 QY 41 LysPProSerGlyAspAspArgTyrAsnGln-PheHisIleIleArg 54
 Db 141 AATCACTCTGGTATGATGATGATACATCAAGTTTCACTAATAGG 183
 RESULT 10
 ID AAD10198 standard; cDNA; 1542 BP.
 XX
 AC AAD10198;
 XX
 DT 24-SEP-2001 (first entry)
 XX

DE Mouse interleukin-1 (IL-1) receptor-associated kinase (IRAK)-4 cDNA.
 XX
 KW Mouse; interleukin-1 receptor associated kinase-4; IRAK-4; cytosolic;
 KW IL; antibacterial; antiinflammatory; ophthalmological; vasotropic; OPD;
 KW chronic obstructive pulmonary disease; neuroprotective; asthma; ARDS;
 KW adult respiratory disease syndrome; pulmonary fibrosis; chronic cough
 KW interstitial lung disease; allergic rhinitis; transplant rejection;
 KW autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;
 KW multiple sclerosis; diabetes; cancer; solid tumour; lymphoma; stroke;
 KW cardiovascular disease; atherosclerosis; neurodegenerative disease;
 KW sepsis; osteoarthritis; rash; osteoporosis; psoriasis; dermatitis;
 KW inflammatory bowel disease; gout; Crohn's disease; ulcerative colitis;
 KW Behcet's syndrome; ankylosing spondylitis; IL-1 receptor/Toll receptor;
 KW sarcoidosis; transgenic animal; ss.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT 1..1542
 FT CDS /*tag= a
 FT /product= "Mouse IRAK-4"
 FT
 XX WO200151641-A1.
 XX 19-JUL-2001.
 XX 12-JAN-2001; 2001WO-US01171.
 XX 13-JAN-2000; 2000US-0176395.
 XX (TULIA-) TULARIK INC.
 XX Wesche H, Li S;
 XX WPI; 2001-451860/48.
 DR P-PSDB; AAE05399.
 DR
 XX Novel human interleukin-1 receptor associated kinase polypeptide,
 PT useful for identifying modulators of the polypeptide for treating gout,
 PT asthma, allergic rhinitis, multiple sclerosis and skin cancer
 PT
 XX Claim 26; Fig 4; 89pp; English.
 XX
 CC The present sequence is a cDNA encoding mouse interleukin (IL)-1 receptor
 CC associated kinase (IRAK)-4. IRAK associate with activated IL-1, IL-18
 CC and other receptors and act to transduce signals originating from the
 CC activated receptors, ultimately leading to a variety of downstream
 CC effects such as nuclear factor (NF)-kappaB activation. The IRAK-4
 CC inhibitors are useful for treating inflammatory diseases such as
 CC pulmonary diseases and diseases of the airway (e.g., adult respiratory
 CC disease syndrome (ARDS), chronic obstructive pulmonary disease (OPD),
 CC pulmonary fibrosis, interstitial lung disease, asthma, chronic cough
 CC or allergic rhinitis), transplant rejection, autoimmune diseases (e.g.,
 CC rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis
 CC or diabetes), cancer (e.g., solid tumours, skin cancer or lymphoma),
 CC cardiovascular diseases (e.g., stroke and atherosclerosis), diseases
 CC of the central nervous system (e.g., neurodegenerative disease), CD14
 CC mediated sepsis, non-CD14 mediated sepsis, osteoarthritis, osteoporosis,
 CC psoriasis, diseases of the skin (e.g., rash, contact dermatitis, atopic
 CC dermatitis), inflammatory bowel disease (e.g., Crohn's disease and
 CC ulcerative colitis), Behcet's syndrome, ankylosing spondylitis, gout,
 CC sarcoidosis and ophthalmic diseases and conditions. The inhibitors of
 CC IRAK-4 activity or expression are used to inhibit signal transduction
 CC resulting from the activation of an interleukin-1 receptor (IL-1R)/Toll
 CC receptor in a cell. They also inhibit the activation of a transcription
 CC factor that activates NFkappaB in the cell. IRAK-4 is used to create a
 CC nonhuman transgenic animal which is useful for testing the function of
 CC IRAK-4 in vivo, to generate models for the study of inflammatory
 CC disorders and conditions and for the development of potential treatments
 CC for IRAK-4-related inflammatory diseases and conditions. IRAK-4 sequences
 CC are also used in gene therapy and in antisense therapy.
 XX
 SQ Sequence 1542 BP; 421 A; 392 C; 423 G; 306 T; 0 other;

PN WO200009552-A1.
 XX 24-FEB-2000.
 PD
 XX 13-AUG-1999; 99WO-US18298.
 PR 14-AUG-1998; 98US-0096622.
 PR 17-AUG-1998; 98US-0096815.
 PR 04-SEP-1998; 98US-0099229.
 PR 23-OCT-1998; 98US-0105368.
 PR 08-JAN-1999; 99US-0115234.
 PR 12-FEB-1999; 99US-0119931.
 PR 18-FEB-1999; 99US-0120575.
 PR 30-APR-1999; 99US-0132020.
 PR 11-AUG-1999; 99US-0096622.
 XX
 PA (GENY) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavalie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Agostino MJ, Steinger RJ, Spaulding V;
 PI Wong GG, Clark HF, Fechel K;
 XX
 DR WPI: 2000-205979/18.
 DR P-PSDB; AAY94963.
 XX
 PT New polynucleotides encoding secreted proteins, which may have e.g.
 PT nutritional, chemokine, immune stimulating or suppressing,
 PT hematopoiesis regulating, tissue growth, activin/inhibin
 PT antiinflammatory or tumor inhibition activity
 XX
 PS Claim 140; Page 590-592; 641pp; English.
 XX
 CC AAA16618 to AAA16697 encode the human secreted proteins given in
 CC AAY94968 to AAY94980, isolated from human adult brain, adult thyroid,
 CC adult retina, foetal carcinoma, adult blood, adult neural, foetal kidney,
 CC adult placenta, adult testis, whole embryo, adult cartilage, kidney,
 CC foetal brain, adult thymus, foetal placenta, adult uterus, adult tumour,
 CC and adult bladder, cDNA libraries. The polynucleotides and proteins are
 CC predicted to have biological activities which would make them suitable
 CC for treating, preventing or ameliorating medical conditions in humans
 CC and animals. The polynucleotides can be used as markers for tissues in
 CC which the protein is preferentially expressed, as molecular weight
 CC markers on Southern gels, and as chromosome markers or tags to identify
 CC chromosomes or to map gene positions. The proteins can be used in the
 CC treatment of immune deficiencies and disorders, such as severe combined
 CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other
 CC infections. These infections include human immunodeficiency virus (HIV),
 CC hepatitis, herpesviruses, mycobacteria, leishmania spp., malaria and
 CC candidiasis. The proteins can be used to treat autoimmune disorders such
 CC as connective tissue disease, multiple sclerosis, systemic lupus
 CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation,
 CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent
 CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and
 CC autoimmune inflammatory eye disease. The proteins can also be used to
 CC treat allergic conditions, such as asthma. AAA16698 to AAA16774 represent
 CC probes for the human secreted proteins from the present invention.
 XX
 SQ Sequence 5170 BP; 1372 A; 1257 C; 1174 G; 1367 T; 0 other;
 Alignment Scores:
 Pred. No.: 19.8 Length: 5170
 Score: 68.00 Matches: 14
 Percent Similarity: 46.34% Conservative: 5
 Best Local Similarity: 34.15% Mismatches: 10
 Query Match: 21.05% Indels: 12
 DB: Gaps: 1
 US-10-001-254-26 (1-59) x AAA16683 (1-5170)
 QY 16 ValGlyLeuLeuLeuGlyLeuSerAspPheIleAspProGlnGluGlyTyrPlyGlyLeu 35
 DB 4523 ATAGTGTCTAATTAACAGAGTGTGATTTGTCAGACCTCCCTCAATGAAA----- 4576

QY 36 AlAlValAlaIleuLysLysProSerGlyAspAspArgTyrAsnGlnPheHisIleArgCys 55
 DB 4577 -----GACAGACTTATATACCTGCATTGACATGC 4606
 QY 56 Cys 56
 DB 4607 TGC 4609
 RESULT 13
 ID ABQ81849 standard; DNA; 349980 BP.
 XX
 AC ABQ81849;
 XX
 DT 19-NOV-2002 (first entry)
 XX
 DE Bifidobacterium longum NCC2705 related nucleotide sequence SEQ ID:1105.
 XX
 KW Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
 KW antidiarrhetic; antibacterial; inhibitor of salmonella; detection;
 KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
 KW rotavirus; food composition; pharmaceutical composition; gene; ds.
 XX
 OS Bifidobacterium longum.
 OS Synthetic.
 OS
 PN EP1227152-A1.
 XX
 PD 31-JUL-2002.
 XX
 PF 30-JAN-2001; 2001EP-0102050.
 XX
 PR 30-JAN-2001; 2001EP-0102050.
 XX
 PA (NEST) SOC PROD NESTLE SA.
 XX
 DR WPI: 2002-668397/72.
 XX
 PT Novel polynucleotide comprising Bifidobacterium genome sequence useful
 PT as a probe or primer for detecting and/or identifying Bifidobacterium
 PT longum in a biological sample
 XX
 PS Disclosure; SEQ ID 1105; 80pp; English.
 XX
 CC The present invention describes a polynucleotide (I) comprising a
 CC sequence of a Bifidobacterium genome selected from the nucleotide
 CC sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
 CC least 90% identity or which hybridises with the sequences given in
 CC ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding
 CC a fusion protein, comprising a sequence selected from 1097 sequences
 CC given in ABP65258 to ABP66354 ligated in frame to a polynucleotide
 CC encoding a heterologous polypeptide. (I) has antidiarrhetic and
 CC antibacterial activities, and can be used as an inhibitor of Salmonella.
 CC (I) (which is a probe) is useful for the detection and/or identification
 CC of Bifidobacterium longum in a biological sample. A carrier containing
 CC the lactic acid bacterium Bifidobacterium longum NCC2705 (CNCM I-2618)
 CC can be used for preventing and/or treating diarrhoea brought about by
 CC pathogenic bacteria and/or rotavirus. The carrier is a food composition
 CC selected from milk, yogurt, curd, cheese, fermented milks, milk based
 CC fermented products, ice-creams, fermented cereal based products, milk
 CC based powders, infant formulae, pet food or a pharmaceutical composition
 CC selected from tablets, liquid bacterial suspensions, dried oral
 CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.
 CC (I) is useful in DNA arrays or chips to carry out analysis of the
 CC expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent
 CC Bifidobacterium related nucleotide sequences given in the Sequence
 CC listing from the present invention but not mentioned further within the
 CC specification.
 CC N.B. The sequence data for this patent is not represented in the printed
 CC specification but is based on sequence information supplied by the
 CC European Patent Office.
 XX
 SQ Sequence 349980 BP; 60046 A; 106490 C; 106389 G; 69055 T; 0 other;

Pred. No.: 24.7
 Score: 66.00
 Percent Similarity: 52.73%
 Best Local Similarity: 34.55%
 Query Match: 20.43%
 DB: 23
 Length: 3610
 Matches: 19
 Conservative: 10
 Mismatches: 20
 Indels: 6
 Gaps: 2

US-10-001-254-26 (1-59) x ABL27970 (1-3610)

QY 2 AsnLysProIleThrProSerThrTyrValArgCysLeuAsnVal----- 16
 DB 2331 AATCTCCCATCGATCCATCCAGCTATGTTCTGGCCATCAGCCAGATGCGGGGGCC 2272
 QY 17 GlyLeuIleArgLysLeuSerAspPheIleAspProGlnGluGlyTyrLysLysLeuAla 36
 DB 2271 CAACCTCTTCGCCAGCTGCGAAAGCCGAGATGCGGAGCCGGAAGTTCACAGAGTGCT 2212
 QY 37 ValAlaIleLysLysPro---SerGlyAspAspArgTyrAsnGln 50
 DB 2211 GCGACTTAAACCCACCCCGTTAATGATGATGATATATATCA 2167

Search completed: January 18, 2004, 01:32:43
 Job time : 98.7617 secs

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 18, 2004, 01:16:08 ; Search time 15.3955 Seconds
(without alignment)
1691.512 Million cell updates/sec

Title: US-10-001-254-26
Perfect score: 383
Sequence: 1 MNKPITPSTYVCINVLIR.....KKPSGDDRYNGFHRCSSQN 59

Scoring table:
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODBL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10001254/runat_16012004_152425_19740/app_query.fasta_1.1109
-DB=Issued_Patents_NA -QPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=D1osum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MITLEN=0 -MATHEN=2000000000
-USER=US10001254 @CGN 1.1_95 @runat_16012004_152425_19740 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESUBSTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	290	89.8	833	4 US-09-166-350-10	Sequence 10, Appl
2	61	18.9	580073	4 US-08-545-528D-1	Sequence 1, Appl
3	59.5	18.4	3381	3 US-09-009-119-1	Sequence 1, Appl
4	59.5	18.4	3381	4 US-09-371-507-1	Sequence 1, Appl
5	59.5	18.4	3381	5 PCT-US95-09098-1	Sequence 1, Appl
6	58	18.0	1182	4 US-09-252-991A-4901	Sequence 4901, Ap
7	58	18.0	3482	3 US-09-111-085-3	Sequence 3, Appl
8	58	18.0	202001	4 US-09-734-674-3	Sequence 3, Appl
9	57.5	17.8	1794	3 US-08-123-934A-9	Sequence 9, Appl
10	57.5	17.8	1794	5 PCT-US94-10080-9	Sequence 9, Appl
11	57.5	17.8	3293	1 US-07-923-976-1	Sequence 1, Appl
12	57	17.6	2462	3 US-09-111-085-1	Sequence 1, Appl

13	56.5	17.5	936	3 US-09-461-697-402	Sequence 402, App
14	56.5	17.5	1017	3 US-09-461-697-396	Sequence 396, App
15	56.5	17.5	1128	3 US-09-461-697-394	Sequence 394, App
16	56.5	17.5	1260	3 US-09-461-697-388	Sequence 388, App
17	56.5	17.5	1467	3 US-09-461-697-374	Sequence 374, App
18	56.5	17.5	1473	3 US-09-461-697-372	Sequence 372, App
19	56.5	17.5	2702	3 US-09-461-697-359	Sequence 359, App
20	56	17.3	941	3 US-08-829-613-1	Sequence 1, Appl
21	56	17.3	1623	4 US-09-513-783A-33	Sequence 33, Appl
22	56	17.3	61663	4 US-09-453-702B-62	Sequence 62, Appl
23	55.5	17.2	1029	2 US-08-484-993B-21	Sequence 21, Appl
24	55.5	17.2	1029	2 US-08-484-158B-21	Sequence 21, Appl
25	55.5	17.2	1029	2 US-08-484-596A-21	Sequence 21, Appl
26	55.5	17.2	1029	2 US-08-480-150A-21	Sequence 21, Appl
27	55.5	17.2	1029	3 US-08-458-731-21	Sequence 21, Appl
28	55.5	17.2	1029	3 US-08-149-223A-21	Sequence 21, Appl
29	55.5	17.2	2238	2 US-08-674-887A-7	Sequence 7, Appl
30	55.5	17.2	2238	3 US-08-951-844-7	Sequence 7, Appl
31	55.5	17.2	2238	4 US-09-412-347-7	Sequence 7, Appl
32	55.5	17.2	2515	4 US-09-221-017B-745	Sequence 745, App
33	55	17.0	420	3 US-08-991-890-1	Sequence 1, Appl
34	55	17.0	438	4 US-08-979-847B-128	Sequence 128, App
35	55	17.0	438	4 US-08-979-847B-129	Sequence 129, App
36	55	17.0	438	4 US-08-979-847B-130	Sequence 130, App
37	55	17.0	438	4 US-08-979-847B-131	Sequence 131, App
38	55	17.0	480	3 US-08-991-890-3	Sequence 3, Appl
39	55	17.0	615	2 US-08-482-842B-1	Sequence 1, Appl
40	55	17.0	615	2 US-08-482-842B-3	Sequence 1, Appl
41	55	17.0	615	2 US-09-174-465D-1	Sequence 1, Appl
42	55	17.0	615	4 US-09-599-564A-1	Sequence 1, Appl
43	55	17.0	934	3 US-09-174-465D-4	Sequence 4, Appl
44	55	17.0	934	4 US-09-599-564A-4	Sequence 4, Appl
45	55	17.0	15144	3 US-08-458-434A-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-166-350-10
; Sequence 10, Application US/09166350A
; Patent No. 6440663
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Lloyd
; APPLICANT: Old, Alex
; APPLICANT: Knuth, Alex
; TITLE OF INVENTION: Renal Cancer Associated Antigens and
; TITLE OF INVENTION: Uses Therefor
; FILE REFERENCE: L0461/7051
; CURRENT APPLICATION NUMBER: US/09/166,350A
; EARLIER FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: US 09/166,350
; EARLIER FILING DATE: 1998-10-05
; NUMBER OF SEQ. ID NOS.: 35
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 833
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-166-350-10
Alignment Scores:
Pred. No.: 8,07e-36
Score: 290.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 89.78%
DB: 4
Gaps: 0
US-10-001-254-26 (1-59) x US-09-166-350-10 (1-833)

Length: 833
Matches: 54
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

```
QY      1 MetAsnLyseProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
      50 ATGAACAAACCCATCAACACATGATGCGCTGCTCAATGTTGACTTAATTAGG 109
      21 LysLeuSerAspPheIleAspProGlnGluGlyTrpLysLysLeuAlaValAlaIleLys 40
      110 AAGCTGTGAGATTATTATGATCTCTCAAGAGAGATGAGAGAGTAGCTGATGCTATTAA 169
QY      41 LysProSerGlyAspAspArgTyArgGlnPheHisIleArg 54
      170 AAACCATCTGGTATGATGATACAAATCAGTTTCACATTAAGG 211
RESULT 2
US-08-545-528D-1
; Sequence 1, Application US/08545528D
; Patent No. 6537773
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragment
; Patent No. 6537773
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB193P1
; CURRENT APPLICATION NUMBER: US/08/545,528D
; PRIOR FILING DATE: 1995-10-19
; PRIOR APPLICATION NUMBER: US 08/488,018
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/473,545
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 580073
; TYPE: DNA
; ORGANISM: Mycoplasma genitalium
US-08-545-528D-1
Alignment Scores:
Pred. No.: 3.13e+04 Length: 580073
Score: 61.00 Matches: 12
Percent Similarity: 46.94% Conservative: 11
Best Local Similarity: 24.49% Mismatches: 8
Query Match: 18.89% Indels: 18
Gaps: 1
DB: 4
US-10-001-254-26 (1-59) x US-08-545-528D-1 (1-580073)
QY      11 ValArgCysLeuAsnValGlyLeuIleArgLysLeuSerAspPhe----- 25
      96165 ATTCAATTGTTTAAACGTTTGTCTCATGACAAAGATGGAAGCGCATTAAGCAATTAAGTA 96224
QY      26 -----11LeaPProGlnGluGlyTrp 32
      96225 ATGAACATTGACAGTGCCTCTTTCCAAAAATGTTAACTGAAGAACCAAAACGGGTGG 96284
QY      33 LysLysLeuAlaValAlaIleLysLys 41
      96285 AAGCAATTAATTGACACTAGAAAA 96311
RESULT 3
US-09-009-119-1/c
; Sequence 1, Application US/09009119
; Patent No. 6160206
; GENERAL INFORMATION:
; APPLICANT: SATO, RYO
; APPLICANT: BOYNTON, John E.
; APPLICANT: GILHAM, Nicholas W.
; APPLICANT: HARRIS, Elizabeth H.
; TITLE OF INVENTION: Porphyryn Accumulating-Type Herbicide Resistance Gene
; FILE REFERENCE: Substitute sequence listing
; CURRENT APPLICATION NUMBER: US/09/009,119
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin Ver. 2.0
```

```
; SEQ ID NO 1
; LENGTH: 3381
; TYPE: DNA
; ORGANISM: Chlamydomonas reinhardtii
US-09-009-119-1
Alignment Scores:
Pred. No.: 45.3 Length: 3381
Score: 59.50 Matches: 17
Percent Similarity: 46.77% Conservative: 12
Best Local Similarity: 27.42% Mismatches: 22
Query Match: 18.42% Indels: 11
Gaps: 2
DB: 4
US-10-001-254-26 (1-59) x US-09-009-119-1 (1-3381)
QY      6 ThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArgLysLeuSerAspPhe 25
      2156 AGCCCGGCTAGCCCTTGGCGGTGATGAGATGAGTGGTGAAGGATTTGGGATATTC 2097
      26 11LeaPProGlnGluGlyTrpLysLysLeu----- 35
      2096 ATGACCCGTGAAGTGGCGGTAAAGTGGCGTGGCGGACGAGACAGGCGATGCGGT 2037
      36 AlaValAlaIleLysLeuProSerGlyAspAspArgTyArgGlnPheHisIleArgCys 55
      2036 GCTCGGACACAGCGCTTGGCTTAGTGCCAGTCCCGGTGATGGCTTGCAA---GGGTGC 1980
QY      56 CysSer 57
      1979 TGTTC 1974
RESULT 4
US-09-371-507-1/c
; Sequence 1, Application US/09371507
; Patent No. 634656
; GENERAL INFORMATION:
; APPLICANT: SATO, RYO
; APPLICANT: BOYNTON, John E.
; APPLICANT: GILHAM, Nicholas W.
; APPLICANT: HARRIS, Elizabeth H.
; TITLE OF INVENTION: Porphyryn Accumulating-Type Herbicide Resistance Gene
; FILE REFERENCE: substitute sequence listing
; CURRENT APPLICATION NUMBER: US/09/371,507
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 09/009,119
; PRIOR FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3381
; TYPE: DNA
; ORGANISM: Chlamydomonas reinhardtii
US-09-371-507-1
Alignment Scores:
Pred. No.: 45.3 Length: 3381
Score: 59.50 Matches: 17
Percent Similarity: 46.77% Conservative: 12
Best Local Similarity: 27.42% Mismatches: 22
Query Match: 18.42% Indels: 11
Gaps: 2
DB: 4
US-10-001-254-26 (1-59) x US-09-371-507-1 (1-3381)
QY      6 ThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArgLysLeuSerAspPhe 25
      2156 AGCCCGGCTAGCCCTTGGCGGTGATGAGATGAGTGGTGAAGGATTTGGGATATTC 2097
      26 11LeaPProGlnGluGlyTrpLysLysLeu----- 35
      2096 ATGACCCGTGAAGTGGCGGTAAAGTGGCGTGGCGGACGAGACAGGCGATGCGGT 2037
```

Qy 36 AlavAlaIlelyslYsPserGlyAspAspArgTyrAsnGlnPheHisIleArgCys 55
Db 2036 GCCTGGCAGACAGCTTGCGCTTAGTGCCAGTCCCGCTGGATGGGCTTGCAA---GGGTGC 1980
Qy 56 CysSer 57
Db 1979 TGTTC A 1974

RESULT 5
PCT-US95-09098-1/c
; Sequence 1, Application PC/TUS9509098
; GENERAL INFORMATION:
; APPLICANT: Sato, Ryo
; APPLICANT: Boynton, John E.
; APPLICANT: Gillham, Nicholas W.
; APPLICANT: Harris, Elizabeth H.
; TITLE OF INVENTION: Porphyrin-Accumulating Type Herbicide
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESS: Birch, Stewart, Kolasch & Birch, LLP
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09098
; FILING DATE: 20-JUL-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 2185-110P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3383 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Chlamydomonas reinhardtii
; STRAIN: RS-3
; PCT-US95-09098-1

Alignment Scores:
Pred. NO.: 45.4 Length: 3383
Score: 59.50 Matches: 17
Percent Similarity: 46.77% Conservative: 12
Best Local Similarity: 27.42% Mismatches: 22
Query Match: 18.42% Indels: 11
DB: Gaps: 2

US-10-001-254-26 (1-59) x PCT-US95-09098-1 (1-3383)

Qy 6 ThrProSerThrTyrValArgCysAlaLeuAsnValGlyLeuIleArgIleuSerAspPhe 25
Db 2157 AGCCCGGCTACCCCTTGCGGTGCTTAGTGCGATGTGGGTGAGTGCAATTGGGAAATATC 2098
Qy 26 IleAspProGlnGluGlyTyrIlePheLysLeu----- 35

Db 2097 ATGACCGGTGAAGTGGCGGTGAAGTGGCGTGGCGGGAGACAGGGCAATGTGGCT 2038
Qy 36 AlavAlaIlelyslYsPserGlyAspAspArgTyrAsnGlnPheHisIleArgCys 55
Db 2037 GCCTGGCAGACAGCTTGCGCTTAGTGCCAGTCCCGCTGGATGGGCTTGCAA---GGGTGC 1981
Qy 56 CysSer 57
Db 1980 TGTTC A 1975

RESULT 6
US-09-252-991A-4901/c
; Sequence 4901, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4901
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4901

Alignment Scores:
Pred. NO.: 18.2 Length: 1182
Score: 58.00 Matches: 9
Percent Similarity: 57.14% Conservative: 7
Best Local Similarity: 32.14% Mismatches: 12
Query Match: 17.96% Indels: 0
DB: Gaps: 0

US-10-001-254-26 (1-59) x US-09-252-991A-4901 (1-1182)

Qy 29 GlnGluGlyTyrPheLysLeuAlaValAlaIlelyslYsPserGlyAspAspArgTyr 48
Db 89 AAACGGCCCTGGAGAGCTTGGCGTCTTGACGTCGCGCCCTGGAGGACAACTGG 30
Qy 49 AsnGlnPheHisIleArgCysCys 56
Db 29 TCCATGATCATCCGGAAGTGTCTGC 6

RESULT 7
US-09-111-085-3/c
; Sequence 3, Application US/09111085
; Patent No. 6100034
; GENERAL INFORMATION:
; APPLICANT: Stoye, Jonathan P
; APPLICANT: Weiss, Robin A
; TITLE OF INVENTION: Detection of retroviral subtypes based upon envelope
; FILE REFERENCE: 4238/75168
; CURRENT APPLICATION NUMBER: US/09/111,085
; CURRENT FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: GB 9710154.7
; EARLIER FILING DATE: 1997-05-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 3482
; TYPE: DNA
; ORGANISM: Porcine retrovirus
US-09-111-085-3

Alignment Scores:

Pred. No.: 80.8 Length: 3482
Score: 58.00 Matches: 18
Percent Similarity: 42.59% Conservative: 5
Best Local Similarity: 33.33% Mismatches: 23
Query Match: 17.96% Indels: 8
DB: 3 Gaps: 3

US-10-001-254-26 (1-59) x US-09-111-085-3 (1-3482)

QY 11 ValArgCysLeuAsnValGlyLeuIleArgLysLeuSerAspPheIleAsp----- 27
DB 1017 GTTAGAGTAGAAGACCGAGCGAGCGAGCTTAAGGGATTCAGCTCTTTTCGGC 958
QY 28 -----ProGlnGluGly---TTPlyLysLeuAlaValAlaIleLysLysProSerGly 44
DB 957 TCTCCACCCCGAGTCGGAGAGTGCGCCAGCTTAAGCTGGAGATGCATGCCAGTGGA 898

QY 45 AspAspArgTyrAsnGlnPheHisIleArgCysCysSerGln 58
DB 897 GATTCC-----TTCCAGCTTTCACAGCCGTTGGTGCTCA 862

RESULT 8
US-09-734-674-3/c
Sequence 3, Application US/09734674
Patent No. 6498022
GENERAL INFORMATION:
APPLICANT: WEI, Ming-Hui et al
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
FILE REFERENCE: C1001018
CURRENT APPLICATION NUMBER: US/09/734,674
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ. ID NOS: 4
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ. ID NO 3
LENGTH: 202001
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(202001)
OTHER INFORMATION: n = A,T,C or G
US-09-734-674-3

Alignment Scores:
Pred. No.: 2.15e+04 Length: 202001
Score: 58.00 Matches: 13
Percent Similarity: 62.96% Conservative: 4
Best Local Similarity: 48.15% Mismatches: 10
Query Match: 17.96% Indels: 0
DB: 4 Gaps: 0

US-10-001-254-26 (1-59) x US-09-734-674-3 (1-202001)

QY 29 GlnGlnGluGlyTTPlyLysLysLeuAlaValAlaIleLysLysProSerGlyAspAspArgTyr 48
DB 91623 AAAAGAGGTACAAAATAATTTGCTGTACTTAATAATACACTAAGGGGAAAAGATATAC 91564

QY 49 AsnGlnPheHisIleArgCys 55
DB 91563 CTGCAATTTCTTCTTAAGTGT 91543

RESULT 9
US-08-123-934A-9/c
Sequence 9, Application US/08123934A
Patent No. 6291206
GENERAL INFORMATION:
APPLICANT: WOZNEY, John
APPLICANT: CELESTE, Anthony J.
APPLICANT: THIES, R. Scott
APPLICANT: YAMAJI, No. 6291206oru
TITLE OF INVENTION: RECEPTOR PROTEINS

NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute Inc. - Legal Affairs
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/123,934A
FILING DATE: 17-SEP-1993
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: LAZAR, Steven R
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5203
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 876 1170
TELEFAX: 617 876 5851
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1794 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: W-120
FEATURE:
NAME/KEY: CDS
LOCATION: 83..1591
US-08-123-934A-9

Alignment Scores:
Pred. No.: 38.8 Length: 1794
Score: 57.50 Matches: 18
Percent Similarity: 43.86% Conservative: 7
Best Local Similarity: 31.58% Mismatches: 29
Query Match: 17.80% Indels: 3
DB: 3 Gaps: 1

US-10-001-254-26 (1-59) x US-08-123-934A-9 (1-1794)

QY 5 IleThrProSerThrTyrValArgCysLeuAsnValGlyLeuIleArgLysLeuSerAsp 24
DB 1587 TTGATCCCTTCTCGTGGCTGAGTTGTACAAATGTTTTTTTATTCGCAAGCTGTACAC 1528

QY 25 PheIleAspPro-----GlnGlnGlyTTPlyLysLysLeuAlaValAlaIleLysLys 41
DB 1527 CTGTGCTGCTCATTTGGCATACAGCATTTCTCATTAATTTTTCATCTCAAGGCC 1468

QY 42 ProSerGlyAspAspArgTyrAsnGlnPheHisIleArgCysCysSerGln 58
DB 1467 TCACAGCTCTGCCATCTGTTGGAAATATTGGCCTTAACCTCTGTTCGCA 1417

RESULT 10
PCT-US94-10080-9/c
Sequence 9, Application PC/TUS9410080
GENERAL INFORMATION:
APPLICANT: GENETICS INSTITUTE, INC.
TITLE OF INVENTION: RECEPTOR PROTEINS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute Inc. - Legal Affairs
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA

ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10080
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,934
FILING DATE: 17-SEP-1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: LAZAR, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5203-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8260
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1794 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: W-120
FEATURE:
NAME/KEY: CDS
LOCATION: 83..1591
PCT-US94-10080-9

Alignment Scores:
Pred. No.: 38.8 Length: 1794
Score: 57.50 Matches: 18
Percent Similarity: 43.86% Conservative: 7
Best Local Similarity: 31.58% Mismatches: 29
Query Match: 17.80% Indels: 3
DB: 5 Gaps: 1

US-10-001-254-26 (1-59) x PCT-US94-10080-9 (1-1794)

QY 5 IleHrProSerThrTyrValArgCysLeuAsnValGlyLeuIleArgGlyLeuSerAsp 24
DB 1587 TTGATGCTTCCTGCTGCTGAGTGTGACATGTTTAAATTCGCAAGCTGTCAGC 1528
QY 25 PheIleAspPro-----GlnGluGlyTrrpLysLysLeuAlaValAlaIleLysLys 41
DB 1527 CTTCCTGCTCATTCGACATACACATCTCTCATTAATTTCAGCATCACTCTCAAGGCC 1468
QY 42 ProSerGlyAspAspArgTyrAsnGlnPheHisIleArgCysCysSerGln 58
DB 1467 TCACAGCTTCGCGCATCTGTTGGAATATTGGCCTTAACCTTCGTCGCA 1417

RESULT 11
US-07-923-976-1/c
Sequence 1, Application US/07923976
Patent No. 5574136
GENERAL INFORMATION:
APPLICANT: Nagata, Shigekazu
APPLICANT: Fukunaga, Rikio
TITLE OF INVENTION: DNA Encoding Granulocyte-
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones, Tullar & Cooper, P.C.
STREET: P.O. Box 2266 Eads Station
CITY: Arlington
STATE: Virginia
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/923,976
FILING DATE: 19920922
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 74539/1990
FILING DATE: 23-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 176629/1990
FILING DATE: 03-JUL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP91/00375
FILING DATE: 22-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hellwege, James W.
REGISTRATION NUMBER: 28,808
REFERENCE/DOCKET NUMBER: 514853
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-415-1508
TELEFAX: 703-415-1500
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3293 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 180..2690
US-07-923-976-1

Alignment Scores:
Pred. No.: 89.6 Length: 3293
Score: 57.50 Matches: 19
Percent Similarity: 35.82% Conservative: 5
Best Local Similarity: 28.36% Mismatches: 14
Query Match: 17.80% Indels: 29
DB: 1 Gaps: 3

US-10-001-254-26 (1-59) x US-07-923-976-1 (1-3293)

QY 12 ArgCysLeuAsnValGlyLeuIleArgLys----- 21
DB 507 CGCTGCTTCCCATGACACTTAAGCAGAGAAGAGCCCTGGCTGACTTCACTGAGGCA 448
QY 22 -----LeuSerAspPheIleAspProGlnGluGly 31
DB 447 GAGTGTAGAGGACCTCTGGGTCCCATCAGCAGATGATGCTGTGTCGCCGTTGGA 388
QY 32 TrpLysLysLeuAlaValAlaIleLysLys---ProSerGlyAspAspArgTyrAsnGln 50
DB 387 TGGGCTCATCTTCGACATCTCCATAGATCTTGGCTGTGTCGA----- 343
QY 51 PheHisIleArgCysCysSer 57
DB 342 -----GTTTGTGCACT 331

RESULT 12
US-09-111-085-1/c
Sequence 1, Application US/09111085
Patent No. 6100034
GENERAL INFORMATION:
APPLICANT: Stoye, Jonathan P
APPLICANT: Weiss, Robin A
TITLE OF INVENTION: Detection of retroviral subtypes based upon envelope
FILE REFERENCE: 4236/75168

```

1 CURRENT APPLICATION NUMBER: US/09/111,085
2
3 CURRENT FILING DATE: 1998-07-07
4
5 EARLIER APPLICATION NUMBER: GB 9710154.7
6
7 EARLIER FILING DATE: 1997-05-16
8
9 NUMBER OF SEQ ID NOS: 16
10
11 SOFTWARE: PatentIn Ver. 2.0
12
13 SEQ ID NO 1
14
15 LENGTH: 2462
16
17 TYPE: DNA
18
19 ORGANISM: Porcine retrovirus
20
21 US-09-111-085-1

```

Alignment Scores:	
pred. No.:	71.8
Score:	57.00
Percent Similarity:	39.34%
Best Local Similarity:	31.15%
Query Match:	17.65%
DB:	3
Length:	2466
Matches:	19
Conservative:	5
Mismatches:	29
Indels:	8
Gaps:	3

US-10-001-254-26 (1-59) X US-09-111-085-1 (1-2462)

Oy		ProIleHxP ^{roSer} HrThrValAlaGySLeuAnVAlGlyLeuIlaArgVLeuSer	23
Dd	338	CCATTAACCTTGAGGACTTATTGCACAGAGTAAGGACCACGCCGTGAGGCCGACCTTAAG	279
Oy	24	AspPheIleaSp-----ProGInglugLy---TrpLysLeuLeuAlaVal	37
Dd	278	GCGATTTTCAGCTCTTTTCGGCTTTCACACCCGAAATCGGAGGTGGCCGCCGCTTAAGGTG	219
Oy	38	AlaIleLysLysP ^{roSer} GrcIYaaPaPaParGTYraGInPheHisIlaArgCysCySeSer	57
Dd	218	GGATGCATGATCATCAGGTGGGAGATTC-----TTCGACTTTCAACAGCCGTTGGTGNGGT	165
Oy		S8 Gln S8	
Dd	164	CNA 162	

```

RESULT 13
; Sequence 402, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Putnam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 402
; LENGTH: 936
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-402

```

Alignment Scores:		
Pred. No.:	22.7	Length: 936
Score:	56.50	Matches: 21
Percent Similarity:	31.46%	Conservative: 7
Best Local Similarity:	23.60%	Mismatches: 26
Query Match:	17.49%	Indels: 35
DB:	3	Gaps: 2

```

Qy      MetAsnIleProIleThrProSerThrTyrValArgCysLeuAsn----- 15
        |||||
        |||
        ::|
Db      1 ATGAATAAAGCTCTACTACTCTCTTTGGGGTGCATGTGGAAATGACCATGTTAATCTA 60
Qy      15 ----- 15
Db      61 CCCATCATGAGACATTTCTTAAGTAAAGCGGCTCTCTCTTGAGCTACCACTTTAA 120
Qy      16 -----ValGlyLeuIleArgLysLeuSerAspPheIleAspPro 28
        |||||
        ::|||
Db      121 GGGAGAGACACCACCGAATCTCTGGAACTTTCTGAATTGGTTGTCTCATTTTGGCAACCT 180
Qy      29 GInGluGlyTrpLysLeuAlaValAlaIleLysLysProSerGlyAspAspArgTyr 48
        |||||
        ::|||
Db      181 ACCGAAGCACTGATTAATGTTCTGAAATACCTCCCAAAA-----AAGATCGAGAA 231
Qy      49 AsnGlnPheHisIleArgCysCysSer 57
        |||||
        ::|||
Db      232 AATGTTAAGAACTTCTGCTGCTGTTCC 258

```

```

RESULT 14
US-09-461-697-396
; Sequence 396, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 396
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-396

```

Alignment Scores:	
Pred. NO.:	25.4
Score:	56.50
Percent Similarity:	31.44%
Best Local Similarity:	23.60%
Query Match:	17.49%
DB:	3
Gaps:	2
Length:	1017
Matches:	21
Conservative:	7
Mismatches:	26
Indels:	35
Gaps:	2

```

US-10-001-254-26 (1-59) x US-09-461-697-396 (1-1017)

QY      1 MetAsnLysProIleThrProSerThryrValArgCysLeuAn-----15
          |||||:::|||||
Db       82 ATGAATAAAGCTTACTACTCTCTTTGGGCGTCATGTGGAAATGACATGTTAATCTA 141
          - - - - -
QY      15 -----15

Db       142 CCCATCATGAGACATTCTTAAGTAAAGCGGCTCTCTCTTGAGACTACCAGTCTAAA 201
          16 -----ValGIyLeuIlleaArgLyIsLeuSeraSpheIIeaSPPro 28
          :::|||:::|||:::|||
Db       202 GGGAGGAGACACCCCGAATCCTGGAGACTTCGCAATTGGTGTCTCATTTTGGCCAACCT 261
          29 GlnGlucLIyTrpLysLysLeuAlaValAlaIleLysLysPProSerGIyaaSPasPARtyr 48
          |||||:::|||
Db       262 ACCGAGCACCTAGATATATGTTCTGAATAATCCCTCCCAAAA-----AAGAGTCAGAA 312

```

```

QY      49 AbglnPheHisIleargCysCysSer 57
          |||  |||  |||  |||  |||
Db      313 AATGTTAAGGAAGCTCTGTGTTCC 339

```

RESULT 15

```

US-09-461-697-394
: Sequence 394, Application US/09461697
: Patent No. 627974
:
: GENERAL INFORMATION:
:   APPLICANT: COGENET NEUROSCIENCE, Inc.
:   APPLICANT: Io, Donald C.
:   APPLICANT: Barney, Shawn
:   APPLICANT: Thomas, Mary Beth
:   APPLICANT: Portbury, Stuart D.
:   APPLICANT: Puranam, Kascluri
:   APPLICANT: Katz, Lawrence C.
:   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
:   TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
:   TITLE OF INVENTION: CELL DEATH
:   FILE REFERENCE: 10001-005-999
:   CURRENT APPLICATION NUMBER: US/09/461,697
:   CURRENT FILING DATE: 1999-12-14
:   NUMBER OF SEQ ID NOS: 466
:   SOFTWARE: FastSeq for Windows Version 4.0
:   SEQ ID NO 394
:   LENGTH: 1128
:   TYPE: DNA
:   ORGANISM: Homo sapiens
US-09-461-697-394

```

Alignment Scores:

Pred. No.:	29.3	Length:	1128
Score:	56.50	Matches:	21
Percent Similarity:	31.46%	Conservative:	7
Best Local Similarity:	23.60%	Mismatches:	26
Query Match:	17.45%	Indels:	35
DB:	3	Gaps:	2

US-10-001-254-26 (1-59) x US-09-461-697-394 (1-1128)

Qy	1	MeAsnIyLeProllethProserThrTyValAlaGysIleuAsn-----	15
		:::	
Db	193	ATGAATAAAGCTTACTACTCTCTTTGGGGTGTCTATGTGAAATGACCATGTAATCTA	255
Qy	15	-----	15
Db	253	CCCATCATGGAGACATTCTTAAAGTAAGCGCGTCTCTCTTGAGAGTACCATTTCTAAA	312
Qy	16	-----ValGlyIleuIleAlaGlyIleuSerAspPheIleAspPro	28
		:::	
Db	313	GGGAGGAGACACCAACCGAATCTCGGGACTTGTGAAATGGGTGTCTCATTTTGGCCAACTCT	372
Qy	29	GIngluGIlyTrpIlysIleuAlaValAlaIleIlyIlysProSerGIlyAspAspArgTyr	48
		:::	
Db	373	ACCGAAGCACTAGATATAGTTCTCGAAATATACCTCCCAAAA-----AAGATATCGAAA	423
Qy	49	AsnGlnPheHisIleAlaGyCysCysSer	57
		:::	
Db	424	AATGTTAAGCAACTTCTGCTGTTCC	450

Search completed: January 18, 2004, 06:11:39
Job time : 73.3955 secs

BEST AVAILABLE COPY

No.	score	Match	Length	DB	ID	Description
-----	-------	-------	--------	----	----	-------------

SUMMARIES

C	1	323	100.0	211	15	US-10-001-224-25	Sequence 25, Appl
C	2	290	89.8	1383	11	US-09-759-595-2	Sequence 2, Appl
C	3	290	89.8	1383	15	US-10-001-224-15	Sequence 15, Appl
C	4	290	89.8	2817	11	US-09-966-451-9	Sequence 3, Appl
C	5	290	89.8	2817	15	US-10-001-224-27	Sequence 27, Appl
C	6	290	89.8	31000	11	US-09-966-451-10	Sequence 10, Appl
C	7	277	85.9	5001	9	US-09-833-799-149	Sequence 149, App
C	8	271	83.9	1542	11	US-09-759-595-4	Sequence 4, Appl
C	9	247	76.5	294	15	US-10-001-224-5	Sequence 5, Appl
C	10	73	22.6	497	13	US-10-027-632-50723	Sequence 50723, A
C	11	73	22.6	497	13	US-10-027-632-50724	Sequence 50724, A
C	12	73	22.6	497	14	US-10-027-632-50723	Sequence 50723, A
C	13	73	22.6	499	14	US-10-027-632-50724	Sequence 50724, A
C	14	73	22.6	499	13	US-10-027-632-12567	Sequence 72567, A
C	15	73	22.6	499	13	US-10-027-632-12568	Sequence 72568, A
C	16	73	22.6	499	14	US-10-027-632-12567	Sequence 72567, A
C	17	73	22.6	499	14	US-10-027-632-12568	Sequence 72568, A
C	18	68	21.1	5170	11	US-09-374-046A-131	Sequence 131, App
C	19	65.5	20.3	4151	9	US-09-764-853-220	Sequence 230, App
C	20	64	19.8	2388	13	US-09-870-133-3	Sequence 3, Appl
C	21	64	19.8	2388	15	US-10-160-501-6	Sequence 6, Appl
C	22	64	19.8	2583	10	US-09-976-674-6	Sequence 6, Appl
C	23	64	19.8	3338	13	US-09-870-133-1	Sequence 1, Appl
C	24	64	19.8	3338	15	US-10-160-501-6	Sequence 4, Appl
C	25	63	19.5	665	13	US-10-027-632-156803	Sequence 156803, A
C	26	63	19.5	665	13	US-10-027-632-156804	Sequence 156804, A
C	27	63	19.5	665	13	US-10-027-632-156805	Sequence 156805, A
C	28	63	19.5	665	14	US-10-027-632-156803	Sequence 156803, A
C	29	63	19.5	665	14	US-10-027-632-156804	Sequence 156804, A
C	30	63	19.5	665	14	US-10-027-632-156805	Sequence 156805, A
C	31	62	19.2	260	12	US-10-359-493-30409	Sequence 30409, A
C	32	61.5	19.2	1889	12	US-10-359-493-29885	Sequence 29885, A
C	33	61.5	19.2	304	9	US-09-864-761-18690	Sequence 18690, A
C	34	61.5	19.0	472	9	US-09-864-761-13929	Sequence 13929, A
C	35	61	18.9	491	11	US-09-918-995-30747	Sequence 30747, A
C	36	61	18.9	6109	13	US-09-795-061-1	Sequence 1, Appl
C	37	61	18.9	13894	13	US-09-960-858-15	Sequence 15, Appl
C	38	61	18.9	13894	13	US-09-960-870-15	Sequence 15, Appl
C	39	61	18.9	580073	13	US-10-245-220-1	Sequence 1, Appl
C	40	60.5	18.7	271930	12	US-10-345-072-87	Sequence 87, Appl
C	41	60.5	18.7	271930	15	US-10-135-144-87	Sequence 87, Appl
C	42	59.5	18.4	2361	12	US-10-108-260A-1556	Sequence 1556, App
C	43	59	18.3	34	11	US-09-759-595-5	Sequence 5, Appl
C	44	59	18.3	498	11	US-09-911-504-163	Sequence 163, Appl
C	45	58.5	18.1	976	10	US-09-974-000-1559	Sequence 1559, Ap

```

TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: CDS
LOCATION: (1) ... (177)
US-10-001-254-25

Alignment Scores:
Pred. No.: 4.5e-43
Score: 323.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 15

US-10-001-254-26 (1-59) x US-10-001-254-25 (1-211)

Qy 1 MetAsnlySPrIoleThrProSeThrTYrValaRgCySLeuAsnVaIGlYLeuIlleaRg 20
Db 1 ATGAACAACCCATACACATACATATGTCGCTGCTCAATGTTGACTAATTGG 60

Qy 21 LysLeuSerAppheilleaSProgInglYlTPlYLSLeuAlaValAlaIlleLys 40
Db 61 AAGGTGCAGATTTTATTGATCTCAAGAAGATGGAAGAAGTATGCTGTAGCTATTAA 120

Qy 41 LysProSeTlYsAppaPaqTYrAaNgInPheIlleAlleaRgCyCysSeRGIaSn 59
Db 121 AAACCACTGCTGATAGTATGATACATCAATGTTCACTAATGAATGCTGTTCCCAAAAC 177

RESULT 2
US-09-759-595-2
Sequence 2, Application US/09759595
Publication No. US2003005916A1
GENERAL INFORMATION:
APPLICANT: Wesche, Holger
APPLICANT: Li, Shun
APPLICANT: Tularik Inc.
TITLE OR INVENTION: IRAK-4: Compositions and Methods of Use
FILE REFERENCE: 018781-003910US
CURRENT APPLICATION NUMBER: US/09/759,595
CURRENT FILING DATE: 2001-01-13
PRIOR APPLICATION NUMBER: US 60/176,395
PRIOR FILING DATE: 2000-01-13
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1383
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human IL-1 receptor-associated kinase 4 (IRAK-4)
OTHER INFORMATION: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: (1) ... (1383)
OTHER INFORMATION: human IRAK-4
US-09-759-595-2

Alignment Scores:
Pred. No.: 1.3e-36
Score: 290.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 89.78%
DB: 11

US-10-001-254-26 (1-59) x US-09-759-595-2 (1-1383)

Qy 1 MetAsnlySPrIoleThrProSeThrTYrValaRgCySLeuAsnVaIGlYLeuIlleaRg 20
Db 1 ATGAACAACCCATACACATACATATGTCGCTGCTCAATGTTGACTAATTGG 60

Qy 21 LysLeuSerAppheilleaSProgInglYlTPlYLSLeuAlaValAlaIlleLys 40

```

```

Db      61 AAGGTGCAATTTTATTTATGATCTCTCAAGAAGATGGAAGAGTTACTGTAAGCTATTAA 120
Qy      41 LysProSeRgLyAspAspArgTYrAsnGlnPheHisIleArg 54
Db      121 AAACCATCTGGTATGATGATAGATACAATCAGTTTCACATAAAGG 162

RESULT 3
US-10-001-254-15
; Sequence 15, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sug Hyung
; APPLICANT: Roth, Wilfred
; APPLICANT: Stemmer-Liwen, Frank
; TITLE OF INVENTION: No. US20030049702A1el Death Domain Proteins
; FILE REFERENCE: P-LJ 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1-1380)
US-10-001-254-15

Alignment Scores:
Pred. No.: 1.3e-36 Length: 1383
Score: 290.00 Matches: 54
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.78% Indels: 0
DB: 15 Gaps: 0

US-10-001-254-26 (1-59) x US-10-001-254-15 (1-1383)

Qy      1 MetAsnLySProIleThrProSerThrTYrValArgCysIleuAsnValGlyIleuIleArg 20
Db      1 ATGAGCAAAACCCATACACATCATATGTCGCGCTCAATGTTGACTTAATTAAAGG 60
Qy      21 LysLeuSerAspPheIleAspProGlnGluGlyTyrLysIleLeuAlaValAlaIleLys 40
Db      61 AAGGTGCAATTTTATTTATGATCTCTCAAGAAGATGGAAGAGTTACTGTAAGCTATTAA 120
Qy      41 LysProSeRgLyAspAspArgTYrAsnGlnPheHisIleArg 54
Db      121 AAACCATCTGGTATGATGATAGATACAATCAGTTTCACATAAAGG 162

RESULT 4
US-09-966-451-3
; Sequence 3, Application US/09966451
; Publication No. US20030087856A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freter
; TITLE OF INVENTION: ANTISENSE MODULATION OF IL-1 RECEPTOR-ASSOCIATED KINASE-4 EXPRESSION
; FILE REFERENCE: RTS-0324
; CURRENT APPLICATION NUMBER: US/09/966,451
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 88
; SEQ ID NO 3

```

LENGTH: 2817
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (50)...(1432)
US-09-966-451-3

Alignment Scores:

Pred. No.: 3,43e-36 Length: 2817
Score: 290.00 Matches: 54
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.78% Indels: 0
Gaps: 0

US-10-001-254-26 (1-59) x US-09-966-451-3 (1-2817)

QY 1 MetAsnlySProlleThrProSerThrTyValArgCysLeuAenValGlyLeuIleArg 20

Db 50 ATGAACAACCCATTAACACCATCAACATATGTGGCTGCTCATGTGGACTAATTAGG 109

QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTrpLysLysLeuAlaValAlaIleLys 40

Db 110 AAGCTGCAGATTATTATGATCTCTCAAGAGATGAGAGAGATTAGCTGTAGCTATTAA 169

QY 41 LysProSerGlyAspAspArgTyAsnGlnPheHisIleArg 54

Db 170 AAACCATCTGCTGATGATGATCAATCAATCAATCAATCAATCAATCAATCAATCA 211

RESULT 5

US-10-001-254-27
Sequence 27, Application US/10001254
Publication No. US20030049702A1

GENERAL INFORMATION:

APPLICANT: Reed, John C.
APPLICANT: Godzik, Adam
APPLICANT: Pawlowski, Krzysztof
APPLICANT: Fiorentino, Loredana
APPLICANT: Lee, Sug Hyung
APPLICANT: Roth, Wilfred
APPLICANT: Steiner-Liewen, Frank
TITLE OF INVENTION: No. US20030049702A1e1 Death Domain Proteins
FILE REFERENCE: P-IJ 5037
CURRENT APPLICATION NUMBER: US/10/001,254
CURRENT FILING DATE: 2001-11-15, 889
PRIOR APPLICATION NUMBER: 60/301,889
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/715,893
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 27
LENGTH: 2817
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: CDS
LOCATION: (50)...(1429)
US-10-001-254-27

Alignment Scores:

Pred. No.: 3.43e-36 Length: 2817
Score: 290.00 Matches: 54
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.78% Indels: 0
Gaps: 0

US-10-001-254-26 (1-59) x US-10-001-254-27 (1-2817)

QY 1 MetAsnlySProlleThrProSerThrTyValArgCysLeuAenValGlyLeuIleArg 20

Db 50 ATGAACAACCCATTAACACCATCAACATATGTGGCTGCTCATGTGGACTAATTAGG 109
QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTrpLysLysLeuAlaValAlaIleLys 40
Db 110 AAGCTGCAGATTATTATGATCTCTCAAGAGATGAGAGAGATTAGCTGTAGCTATTAA 169
QY 41 LysProSerGlyAspAspArgTyAsnGlnPheHisIleArg 54
Db 170 AAACCATCTGCTGATGATGATCAATCAATCAATCAATCAATCAATCAATCAATCA 211

RESULT 6

US-09-966-451-10
Sequence 10, Application US/09966451
Publication No. US20030087856A1

GENERAL INFORMATION:

APPLICANT: C. Frank Bennett
TITLE OF INVENTION: ANTISENSE MODULATION OF IL-1 RECEPTOR-ASSOCIATED KINASE-4 EXPRES
FILE REFERENCE: RTS-0324
CURRENT APPLICATION NUMBER: US/09/966,451
CURRENT FILING DATE: 2001-09-28

NUMBER OF SEQ ID NOS: 88

SEQ ID NO 10

LENGTH: 31000

TYPE: DNA

ORGANISM: Homo sapiens

Alignment Scores:

Pred. No.: 9.03e-35 Length: 31000
Score: 290.00 Matches: 54
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.78% Indels: 0
Gaps: 0

US-10-001-254-26 (1-59) x US-09-966-451-10 (1-31000)

QY 1 MetAsnlySProlleThrProSerThrTyValArgCysLeuAenValGlyLeuIleArg 20

Db 9860 ATGAACAACCCATTAACACCATCAACATATGTGGCTGCTCATGTGGACTAATTAGG 9979

QY 21 LysLeuSerAspPheIleAspProGlnGluGlyTrpLysLysLeuAlaValAlaIleLys 40

Db 9920 AAGCTGCAGATTATTATGATCTCTCAAGAGATGAGAGAGATTAGCTGTAGCTATTAA 9979

QY 41 LysProSerGlyAspAspArgTyAsnGlnPheHisIleArg 54

Db 9980 AAACCATCTGCTGATGATGATCAATCAATCAATCAATCAATCAATCAATCAATCA 10021

RESULT 7

US-09-833-790-149
Sequence 149, Application US/09833790
Patent No. US2002006828A1

GENERAL INFORMATION:

APPLICANT: Lodes, Michael J.
APPLICANT: Wang, Tonglong
APPLICANT: Secrist, Heather
APPLICANT: Mohamath, Raodoh
APPLICANT: Indirias, Carol Y.
TITLE OF INVENTION: PAN, LIQUN
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.512
CURRENT APPLICATION NUMBER: US/09/833,790
CURRENT FILING DATE: 2001-04-11

NUMBER OF SEQ ID NOS: 440

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 149

LENGTH: 501

TYPE: DNA

ORGANISM: Homo sapien

```
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1)...(501)
/ OTHER INFORMATION: n = A,T,C or G
US-09-833-790-149

Alignment Scores:
Pred. No.: 4.17e-35 Length: 501
Score: 277.00 Matches: 54
Percent Similarity: 98.18% Conservative: 0
Best Local Similarity: 98.18% Mismatches: 0
Query Match: 85.76% Indels: 1
DB: 9 Gaps: 0

US-10-001-254-26 (1-59) x US-09-833-790-149 (1-501)

QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
DB 21 ATGAACAAACCCATTAACACCATCAATATGTGCGCTCAATGTGGACTAATTAAG 80
QY 21 LysLeuSerAspPheIleAspProGlnGlyTrpLysLysLeuAlaIleLys 40
DB 81 AAGCTGTCAATTTTATGATCTCAAGAGATGAGAGAGATTAGCTAGCTATTATA 140
QY 41 LysProSerGlyAspAspArgTyraGln-PheHisIleArg 54
DB 141 AAACCATCTGGTATGATATGATCAATCAAGTTTCAATTAAG 183

RESULT 8
US-09-759-595-4
/ Sequence 4, Application US/09759595
/ Publication No. US20030059916A1
/ GENERAL INFORMATION:
/ APPLICANT: Mesche, Holger
/ APPLICANT: Li, Shyun
/ TITLE OF INVENTION: IRAK-4: Compositions and Methods of Use
/ FILE REFERENCE: 018781-003910US
/ CURRENT APPLICATION NUMBER: US/09/759,595
/ CURRENT FILING DATE: 2001-01-13
/ PRIOR APPLICATION NUMBER: US 60/176,395
/ PRIOR FILING DATE: 2000-01-13
/ NUMBER OF SEQ ID NOS: 7
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4
/ LENGTH: 1542
/ TYPE: DNA
/ ORGANISM: Mus sp.
/ FEATURE:
/ OTHER INFORMATION: murine IL-1 receptor-associated kinase 4 (IRAK-4)
/ OTHER INFORMATION: cdna
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (163)..(1542)
/ OTHER INFORMATION: murine IRAK-4
US-09-759-595-4

Alignment Scores:
Pred. No.: 1.81e-33 Length: 1542
Score: 271.00 Matches: 49
Percent Similarity: 98.15% Conservative: 4
Best Local Similarity: 90.74% Mismatches: 1
Query Match: 83.90% Indels: 0
DB: 11 Gaps: 0

US-10-001-254-26 (1-59) x US-09-759-595-4 (1-1542)

QY 1 MetAsnLysProIleThrProSerThrTyValArgCysLeuAsnValGlyLeuIleArg 20
DB 163 ATGAACAAACCCATTAACACCATCAATATGTGCGCTCAATGTGGACTAATTAAG 222
QY 21 LysLeuSerAspPheIleAspProGlnGlyTrpLysLysLeuAlaIleLys 40
```

```
DB 223 AAGCTGTGGAATTTTATGATCTCAAGAGGTGGAAAGAAATTAGAGTATCAAA 282
QY 41 LysProSerGlyAspAspArgTyraGln-PheHisIleArg 54
DB 283 AAGCCGTCCGGCAGCAGCATATCAATCAATGTTCCATTAAG 324

RESULT 9
US-10-001-254-5
/ Sequence 5, Application US/10001254
/ Publication No. US20030049702A1
/ GENERAL INFORMATION:
/ APPLICANT: Reed, John C.
/ APPLICANT: Godzik, Adam
/ APPLICANT: Pawlowski, Krzysztof
/ APPLICANT: Fiorentino, Loredana
/ APPLICANT: Lee, Sung Hyung
/ APPLICANT: Roth, Wilfried
/ APPLICANT: Stenmer-Liwen, Frank
/ TITLE OF INVENTION: No. US20030049702A1 Death Domain Proteins
/ FILE REFERENCE: P-LT 5037
/ CURRENT APPLICATION NUMBER: US/10/001,254
/ CURRENT FILING DATE: 2001-11-15
/ PRIOR APPLICATION NUMBER: 60/301,889
/ PRIOR FILING DATE: 2001-06-29
/ PRIOR APPLICATION NUMBER: 09/715,893
/ PRIOR FILING DATE: 2000-11-17
/ NUMBER OF SEQ ID NOS: 62
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5
/ LENGTH: 294
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)...(294)
US-10-001-254-5

Alignment Scores:
Pred. No.: 1.47e-30 Length: 294
Score: 247.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.47% Indels: 0
DB: 15 Gaps: 0

US-10-001-254-26 (1-59) x US-10-001-254-5 (1-294)

QY 9 ThrTyValArgCysLeuAsnValGlyLeuIleArgLysLeuSerAspPheIleAspPro 28
DB 1 ACATATGTGCGCTGCTCAATGTTGACTAATTAAGAACTGTCAATTTTATGATCCT 60
QY 29 GlnGlnGlyTrpLysLysLeuAlaIleLysLysProSerGlyAspAspArgTyra 48
DB 61 CAAGAGAGATGAGAGAGATTAGCTGTAGCTATTAAAAACCATGTGGTATATGATAC 120
QY 49 AsnGlnPheHisIleArg 54
DB 121 AATCAGTTTCACATTAAG 138

RESULT 10
US-10-027-632-50723/C
/ Sequence 50723, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ POLYMORPHISMS IN THE HUMAN GENOME
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
```

```
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50723
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-50723

Alignment Scores:
Pred. No.: 0.0477 Length: 497
Score: 73.00 Matches: 12
Percent Similarity: 69.57% Conservative: 4
Best Local Similarity: 52.17% Mismatches: 7
Query Match: 22.60% Indels: 0
DB: 13 Gaps: 0

US-10-001-254-26 (1-59) x US-10-027-632-50723 (1-497)

QY 22 LeuserAspPheIleapProGIngluGlyTTrpLysLysLeuAlaValAlaIleLysLys 41
Db 304 CTTCTAGATTGTTGTCCTCCCKACAGAGGTTGGAAGTCTGCTGCAGCCCTGGCCTCT 245
QY 42 ProSerGly 44
Db 244 CCAAGCTGGC 236

RESULT 11
US-10-027-632-50724/c
; Sequence 50724, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50724
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-50724

Alignment Scores:
Pred. No.: 0.0477 Length: 497
Score: 73.00 Matches: 12
Percent Similarity: 69.57% Conservative: 4
Best Local Similarity: 52.17% Mismatches: 7
Query Match: 22.60% Indels: 0
DB: 13 Gaps: 0
```

```
Percent Similarity: 69.57% Conservative: 4
Best Local Similarity: 52.17% Mismatches: 7
Query Match: 22.60% Indels: 0
DB: 13 Gaps: 0

US-10-001-254-26 (1-59) x US-10-027-632-50724 (1-497)

QY 22 LeuserAspPheIleapProGIngluGlyTTrpLysLysLeuAlaValAlaIleLysLys 41
Db 304 CTTCTAGATTGTTGTCCTCCCKACAGAGGTTGGAAGTCTGCTGCAGCCCTGGCCTCT 245
QY 42 ProSerGly 44
Db 244 CCAAGCTGGC 236

RESULT 12
US-10-027-632-50723/c
; Sequence 50723, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50723
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-50723

Alignment Scores:
Pred. No.: 0.0477 Length: 497
Score: 73.00 Matches: 12
Percent Similarity: 69.57% Conservative: 4
Best Local Similarity: 52.17% Mismatches: 7
Query Match: 22.60% Indels: 0
DB: 14 Gaps: 0

US-10-001-254-26 (1-59) x US-10-027-632-50723 (1-497)

QY 22 LeuserAspPheIleapProGIngluGlyTTrpLysLysLeuAlaValAlaIleLysLys 41
Db 304 CTTCTAGATTGTTGTCCTCCCKACAGAGGTTGGAAGTCTGCTGCAGCCCTGGCCTCT 245
QY 42 ProSerGly 44
Db 244 CCAAGCTGGC 236

RESULT 13
US-10-027-632-50724/c
; Sequence 50724, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
```

```

US-10-027-632-72567
Alignment Scores:
Pred. No.: 0.048 Length: 499
Score: 73.00 Matches: 12
Percent Similarity: 69.57% Conservative: 4
Best Local Similarity: 52.17% Mismatches: 7
Query Match: 22.60% Indels: 0
DB: 13 Gaps: 0
US-10-001-254-26 (1-59) x US-10-027-632-72567 (1-499)
QY 22 LeuSerAspPheIleAspProGlnGluGlyTrpLysLysLeuAlaValAlaIleLysLys 41
||| |||||::: ||| ::||| ||||| ||||| |||||
DB 306 CTTCATGATTTCGTGTCCTTACCAAGGTTGGAAGACTGTGGCTGCAGCCTTGCTCA 247
|||:::|
246 CCAAGCTGAC 238

RESULT 15
US-10-027-632-72568/c
Sequence 72568, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027/632
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 72568
LENGTH: 499
TYPE: DNA
ORGANISM: Human
US-10-027-632-72568
Alignment Scores:
Pred. No.: 0.048 Length: 499
Score: 73.00 Matches: 12
Percent Similarity: 69.57% Conservative: 4
Best Local Similarity: 52.17% Mismatches: 7
Query Match: 22.60% Indels: 0
DB: 13 Gaps: 0
US-10-001-254-26 (1-59) x US-10-027-632-72568 (1-499)
QY 22 LeuSerAspPheIleAspProGlnGluGlyTrpLysLysLeuAlaValAlaIleLysLys 41
||| |||||::: ||| ::||| ||||| ||||| |||||
DB 306 CTTCATGATTTCGTGTCCTTACCAAGGTTGGAAGACTGTGGCTGCAGCCTTGCTCA 247
|||:::|
246 CCAAGCTGAC 238

Search completed: January 18, 2004, 09:06:06

```

Tue Jan 20 08:30:19 2004

us-10-001-254-26.rnpb

Page 7

Job time : 93.2609 secs

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 23:09:11 ; Search time 1315 Seconds
(without alignments)
5433.852 Million cell updates/sec

Title: US-10-001-254-5
Perfect score: 294
Sequence: 1 acatagtcgctgcctcaaa.....tgctccagatgctgtccc 294

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
EST:*
1: em_estbda:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estio:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	294	100.0	859	10	BG164491 602342026
2	292.4	99.5	811	10	BG164438 602642772
3	267.2	90.9	719	10	BF696981 602130160
4	262	89.1	402	10	BE482619 168463 BA

5	220	74.8	541	10	BG691069	BG691069 340084 BA
6	205.4	69.9	313	12	BM151935	BM151935 TCBAPE11
7	204.4	69.5	503	14	CA538859	CA538859 C0272B01-
8	204.4	69.5	598	13	BQ552228	BQ552228 HA014C09-
9	204.4	69.5	610	10	BB660378	BB660378 BB660378
10	204.4	69.5	637	10	BB613447	BB613447 BB613447
11	204.4	69.5	638	14	BY721552	BY721552 BY721552
12	204.4	69.5	663	14	BY726858	BY726858 BY726858
13	204.4	69.5	676	10	BB613167	BB613167 BB613167
14	204.4	69.5	1161	11	AK020397	AK020397 Mus muscu
15	204.4	69.5	2461	11	AK028837	AK028837 Mus muscu
16	204.4	69.5	2810	11	AK029028	AK029028 Mus muscu
17	202.8	69.0	575	4	BX522921	BX522921 RZPD Mus
18	198.8	67.6	265	9	AW436511	AW436511 76774 MAR
19	182.4	62.0	507	10	BB666698	BB666698 BB666698
20	181.4	61.7	524	9	AW106160	AW106160 um23h11.y
21	163.2	55.5	453	10	BB660349	BB660349 BB660349
22	147.6	50.2	852	13	BU209111	BU209111 603950834
23	144.4	49.1	600	9	AJ453616	AJ453616 AJ453616
24	144.4	49.1	670	9	AJ447581	AJ447581 AJ447581
25	139.8	47.6	540	9	AL699213	AL699213 DKFZP586K
26	139.6	47.5	505	10	BE482323	BE482323 168064 BA
27	139.4	47.4	629	9	AL647125	AL647125 AL647125
28	137.6	46.8	284	9	AA114228	AA114228 zn75905.x
29	134	45.6	576	10	BF238344	BF238344 601904613
30	125.8	42.8	270	10	BB45841	BB45841 232716 BA
31	123.4	42.0	664	12	BJ035962	BJ035962 BJ035962
32	113.6	38.6	858	10	BF687921	BF687921 60206996
33	107.4	36.5	520	10	BE132064	BE132064 dB41h09.y
34	106.2	36.1	632	14	CA365604	CA365604 640775 NC
35	104.8	35.6	795	13	BQ293310	BQ293310 EST01080
36	104.2	35.4	781	10	BF723681	BF723681 EST00673
37	94.8	32.2	555	9	AW423082	AW423082 f168b07.y
38	94.8	32.2	743	14	CA474136	CA474136 AGENCOURT
39	92.8	31.6	380	12	BI883790	BI883790 fn07c09.y
40	86.4	29.4	490	9	AJ443799	AJ443799 AJ443799
41	79.6	27.1	595	28	AO601806	AO601806 HS 2108 B
42	69	23.5	335	13	BU430802	BU430802 UI-HF-BNO
43	45.6	15.5	873	28	BH150262	BH150262 ENTERF28TR
44	44.2	15.0	318	14	CB940191	CB940191 IPG6X14
45	42.6	14.5	491	9	AI802160	AI802160 lx25a02.x

ALIGNMENTS

RESULT 1
BG164491
LOCUS
DEFINITION
602342026F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4452055 5',
mRNA sequence.
ACCESSION
BG164491
VERSION
BG164491.1 GI:12671194
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 859)
NIH-MGC <http://imgc.ncbi.nih.gov/>.
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
COMMENT
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LHAM10240 row: d column: 08
High quality sequence stop: 634.

FEATURES
source

Location/Qualifiers
1. 859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4452055"
/tissue_type="hypermphroma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pcMV-SPORE6; Site: 1; NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."
BASE COUNT
266 a 168 c 190 g 235 t
ORIGIN

Query Match 100.0%; Score 294; DB 10; Length 859;
Best Local Similarity 100.0%; Pred. No. 1.6e-71;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACATATGCGCGCTGCTCAATGTTGACATTAATGAGAGCTGCAGATTTTATGATCCT 60
DB 59 ACATATGCGCGCTGCTCAATGTTGACATTAATGAGAGCTGCAGATTTTATGATCCT 118
QY 61 CAAGAAGATGAGAGAGTTAGCTAGTATTAATAAACCATCTGATGATAGATAC 120
DB 119 CAAGAAGATGAGAGAGTTAGCTAGTATTAATAAACCATCTGATGATAGATAC 178
QY 121 AATCAGTTTCATTAAGAGATTGAGCATTAATCAAACTGAGAAAAGTCCACTTCT 180
DB 179 AATCAGTTTCATTAAGAGATTGAGCATTAATCAAACTGAGAAAAGTCCACTTCT 238
QY 181 GAATTAAGTTTACATGAGGACACCAAAATGACAGCTTGATCTTGATCTTTTG 240
DB 239 GAATTAAGTTTACATGAGGACACCAAAATGACAGCTTGATCTTGATCTTTTG 298
QY 241 ATCCAAATGAATTTTTCCTGCTGAGAGCTTTTCTCCAGATGCTGTGCC 294
DB 299 ATCCAAATGAATTTTTCCTGCTGAGAGCTTTTCTCCAGATGCTGTGCC 352

RESULT 2
BG616438 811 bp mRNA linear EST 18-APR-2001
LOCUS BG616438
DEFINITION BG616438 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4773760 5',
mrna sequence.

ACCESSION BG616438
VERSION BG616438
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
1 (bases 1 to 811)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLNL at:
http://image.llnl.gov
Plate: L10M1645 row: p column: 17
High quality sequence start: 3
High quality sequence stop: 613.
Location/Qualifiers
1. 811
/organism="Homo sapiens"

FEATURES
source

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4773760"
/tissue_type="embryonal carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_61"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site: 1; SfiI (ggccgctggcc); Site 2: SfiI (ggccatcggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCTAGAGCGCGGCGCGCCGATG-3' and 3' adaptor sequence: 5'-CAGCGCATTTATGCG-3' and 3' adaptor insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."
BASE COUNT
246 a 196 c 164 g 205 t
ORIGIN

Query Match 99.5%; Score 292.4; DB 10; Length 811;
Best Local Similarity 99.7%; Pred. No. 4.4e-71;
Matches 293; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACATATGCGCGCTGCTCAATGTTGACATTAATGAGAGCTGCAGATTTTATGATCCT 60
DB 77 ACATATGCGCGCTGCTCAATGTTGACATTAATGAGAGCTGCAGATTTTATGATCCT 136
QY 61 CAAGAAGATGAGAGAGTTAGCTAGTATTAATAAACCATCTGATGATAGATAC 120
DB 137 CAAGAAGATGAGAGAGTTAGCTAGTATTAATAAACCATCTGATGATAGATAC 196
QY 121 AATCAGTTTCATTAAGAGATTGAGCATTAATCAAACTGAGAAAAGTCCACTTCT 180
DB 197 AATCAGTTTCATTAAGAGATTGAGCATTAATCAAACTGAGAAAAGTCCACTTCT 256
QY 181 GAATTAAGTTTACATGAGGACACCAAAATGACAGCTTGATCTTGATCTTTTG 240
DB 257 GAATTAAGTTTACATGAGGACACCAAAATGACAGCTTGATCTTGATCTTTTG 316
QY 241 ATCCAAATGAATTTTTCCTGCTGAGAGCTTTTCTCCAGATGCTGTGCC 294
DB 317 ATCCAAATGAATTTTTCCTGCTGAGAGCTTTTCTCCAGATGCTGTGCC 370

RESULT 3
BF696981 719 bp mRNA linear EST 22-DEC-2000
LOCUS BF696981
DEFINITION 602130160F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4287014 5',
mrna sequence.

ACCESSION BF696981
VERSION BF696981
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
1 (bases 1 to 719)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLNL at:
http://image.llnl.gov
Plate: L10M124 row: g column: 15
High quality sequence stop: 632.

FEATURES
source

FEATURES

Source

Location/Qualifiers

1. 719
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4287014"
/issue_type="Primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1ib="NIH MGC 56"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgccctggcc); Site_2: SfiI (ggcatttggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCGCATATGAGCC-3' and 3' adaptor sequence: 5'-ATCTAGAGCGCGCGCGCATG-3' (30)BN-3' (where B = A, C, G or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT

216 a 145 c 163 g 195 t

ORIGIN

Query Match 90.9%; Score 267.2; DB 10; Length 719;
Best Local Similarity 98.3%; Pred. No. 4.8e-64;
Matches 291; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 ACATATGCGCGCTCCCAATGTTGACTAATTAGAAAGTCGATGATTTATGATCCT 60
DB 92 ACATATGCGCGCTCCCAATGTTGACTAATTAGAAAGTCGATGATTTATGATCCT 151
QY 61 CAAGAAGATGGAAGAGTTAGTGTAGTCTTAAAAAACCATCTGGATGATAGATAC 120
DB 152 CAAGAAGATGGAAGAGTTAGTGTAGTCTTAAAAAACCATCTGGATGATAGATAC 211
QY 121 AATCAGTTTCAATAGAGAGATTGTAAGCATTTCTCAAACTGGAAGAAAGTCCACTCT 180
DB 212 AATCAGTTTCAATAGAGAGATTGTAAGCATTTCTCAAACTGGAAGAAAGTCCACTCT 271
QY 181 GAATTAATCTGTTGACGCGGCGCACCAAAATGTCAGAGTGTATCTTGGATCTTTT 239
DB 272 GAATTAATCTGTTGACGCGGCGCACCAAAATGTCAGAGTGTATCTTGGATCTTTT 331
QY 240 GATCCAAATGAATTTT-TTGTCTCTCGAGTCTTTTGTCTCCGAGATGCTTCCC 294
DB 332 GATCCAAATGAATTTTGTGCTCTCTCGAGTCTTTTGTCTCCGAGATGCTTCCC 387

RESULT 4

BE482619

LOCUS

168463 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.

DEFINITION

BE482619.1 GI:9602152

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Quackenbush, J.
Analysis of bovine mammary gland EST and functional annotation of the Bos taurus gene index

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Sonstegard TS
USDA, ARS, Beltsville Agricultural Research Center
Belg. 200 Rm 2A, Beltsville, MD 20705, USA
Tel: 301 504 8416

Fax: 301 504 8414
Email: tades@ipri.barc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.

PCR primers
FORWARD: AGGAACAGCATGACCAT
BACKWARD: GTTTCAGTCACGACG
Plate: 11 row: F column: 7
Seq primer: ATTAGGTGACATATAG.

FEATURES

Source

Location/Qualifiers

1. 402
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/issue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI; Library made from pooled mRNA isolated from mammary tissues at eight physiological, developmental, and disease states."

BASE COUNT

114 a 95 c 88 g 105 t

ORIGIN

Query Match 89.1%; Score 262; DB 10; Length 402;
Best Local Similarity 93.2%; Pred. No. 1.1e-62;
Matches 274; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ACATATGCGCGCTCCCAATGTTGACTAATTAGAAAGTCGATGATTTATGATCCT 60
DB 88 ACATATGCGCGCTCCCAATGTTGACTAATTAGAAAGTCGATGATTTATGATCCT 147
QY 61 CAAGAAGATGGAAGAGTTAGTGTAGTCTTAAAAAACCATCTGGATGATAGATAC 120
DB 148 CAAGAAGATGGAAGAGTTAGTGTAGTCTTAAAAAACCATCTGGATGATAGATAC 207
QY 121 AATCAGTTTCAATAGAGAGATTGTAAGCATTTCTCAAACTGGAAGAAAGTCCACTCT 180
DB 208 AATCAGTTTCAATAGAGAGATTGTAAGCATTTCTCAAACTGGAAGAAAGTCCACTCT 267
QY 181 GAATTAATCTGTTGACGCGGCGCACCAAAATGTCAGAGTGTATCTTGGATCTTTT 240
DB 268 GAATTAATCTGTTGACGCGGCGCACCAAAATGTCAGAGTGTATCTTGGATCTTTT 327
QY 241 ATCCAAATGAATTTTGTCTCTCTCGAGTCTTTTGTCTCCGAGATGCTTCCC 294
DB 328 ATCCAAATGAATTTTGTCTCTCTCGAGTCTTTTGTCTCCGAGATGCTTCCC 381

RESULT 5

BG691069

LOCUS

340084 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.

DEFINITION

BG691069.1 GI:13932889

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Quackenbush, J.
Analysis of bovine mammary gland EST and functional annotation of the Bos taurus gene index

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Sonstegard TS
USDA, ARS, Beltsville Agricultural Research Center

Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
Tel: 301 504 8416
Fax: 301 504 8414
Email: cade@lpsl.barc.usda.gov
Single pass sequencing. Bases called and alt. trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.

PCR Primers

FORWARD: AGGAAAGCTATGACCAT
BACKWARD: GTTTCCAGTCACGACG
Plate: 100 row: B column: 15
Seq primer: ATTAGGTGACTATAG.
Location/Qualifiers

FEATURES

source

1. .541
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/issue_type="pooled"
/lab_host="DH10B"
/clone_lib="BARC 580V"
/note="Vector: PCMV SPOR6; Site 1: NotI; Site 2: SalI;
library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."

BASE COUNT 169 a 96 c 119 g 157 t

ORIGIN

Query Match 74.8%; Score 220; DB 10; Length 541;
Best Local Similarity 93.9%; Pred. No. 7.2e-51;
Matches 229; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 51 TATGATCCCTCAAGAGATGAAGAAGTAGGTAGTATTAATAAACCATCTGCTGTA 110
1 TATTGATCTCCAGAAAGATGAGAGAGTAGAGTCGCTATTAATAAACCATCTGCTGA 60
DB 111 TGATGATACAAATCACTTTTCACATPAAGAGATTGAAGCTTCAATCTGAATAAG 170
61 TGATGATACAAATCACTTTTCACATPAAGAGATTGAAGCTTCAATCTGAATAAG 120
QY 171 TCCCATCTTCGAATTAATCTGTTGACCTGGGGACCAACAATTCGACAGTTGATCTTGT 230
121 CCCCAAGGTGAGATTAATCTGTTGACCTGGGGACCAACAATTCGACAGTTGATCTTGT 180
DB 231 GGATCTTTGATCAAAATGAATTTTGGCTCGGAGCTTTGGCTCCAGATGCTGT 290
181 GGATATTTTGGTCCAAATGAGTTTGGCTCCAGCTTTTGGCTACCAATGCTGT 240
QY 291 TCCC 294
DB 241 ACCC 244

RESULT 6 313 bp mRNA linear EST 30-NOV-2001
BM151935
LOCUS TCBAP1B11457 Pediatric pre-B cell acute lymphoblastic leukemia
DEFINITION Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCBAP1145, mRNA
sequence.

ACCESSION BM151935 GI:17175815
VERSION BM151935.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 313)
Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R., Jr.,
Gunaratne, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.

Pediatric leukemia cDNA Sequencing Project (2001)
Unpublished

CONTACT: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine

1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@tccc.org
Seq primer: M13 primer.
Location/Qualifiers

FEATURES

source

1. .313
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TCBAP1145"
/sex="male"
/issue_type="leukopheresis"
/cell_type="pre-B cell"
/dev_stage="pediatric 2 years"
/lab_host="DH10B"
/clone_lib="Pediatric pre-B cell acute lymphoblastic
leukemia Baylor-HGSC project-TCBA"
/note="Vector: lambda pSB; Site 1: BamHI; Site 2: EcoRI;
first strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGACTCGAGCGCGCGAGAGAG (T) VN
3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second strand
was primed with a BamHI-dC primer
[5'AGAGACTCGAGTCGCGCGCGCAATTAATPAT (C) 3']
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P, Westover A, Nishiyama Y, Ohsuni T,
Itoh M, Nagaoka S, Sasaki, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap triapper,
DNA Res 4: 1, 61-6, Feb 28, 1997)"

BASE COUNT 82 a 62 c 81 g 87 t 1 others

ORIGIN

Query Match 69.9%; Score 205.4; DB 12; Length 313;
Best Local Similarity 92.7%; Pred. No. 7e-47;
Matches 215; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 ACATATGTCGCTGCTCAATGTTGACCTAATTAAGAGCTGCAGTTTATATGCTCT 60
77 ACATATGTCGCTGCTCAATGTTGACCTAATTAAGAGCTGCAGTTTATATGCTCT 136
DB 61 CAAGAAGATGAGAGAGTTAGCTGATTAATAAAACATCTGGATGATAGATAC 120
137 CAAGAAGATGAGAGAGTTAGCTGATTAATAAAACATCTGGATGATAGATAC 196
DB 121 AATCAGTTTCACATPAAGAGATTGAAGCATTAATCACTGAGAAAAGTCCACTTCT 180
197 AATCAGTTTCACATPAAGAGATTGAAGCATTAATCACTGAGAAAAGTCCACTTCT 256
QY 181 GAATTACTGTTGACTGCGGACCAACAATTCGACAGTTGGATGCTTGG 232
257 GAATTACTGTTGACTGCGGACCAACAATTCGACAGTTGGATGCTTGG 308
DB

RESULT 7 503 bp mRNA linear EST 19-NOV-2002
CA538859
LOCUS C0272B01-5N NIA Mouse 7.5-dpc Whole Embryo cDNA Library (long) Mus
DEFINITION musculus cDNA clone NIA:C0272B01 IMAGE:30017484 5', mRNA sequence.
ACCESSION CA538859 GI:25080828
VERSION CA538859.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 503)
Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Tanaka, T., Lim, M.K.,
Luo, A. and Ko, M.S.H.

Systematic Analyses of NIA Mouse 7.5-dpc Whole Embryo cDNA Library
TITLE

JOURNAL
COMMENT

(Long)
Unpublished
Other ESTs: C0272B01-3
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: C0272 row: B column: 01
Seq primer: M13 Reverse
High quality sequence stop: 503
POLYA=No.

FEATURES
SOURCE

Location/Qualifiers
1. 593
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:C0272B01-5N"
/db_xref="taxon:10090"
/clone="NIA:C0272B01 IMAGE:30017484"
/tissue_type="whole embryo including extraembryonic
tissues at 7.5-days postcoitum"
/dev_stage="7.5-days postcoitum"
/lab_host="DH10B"
/clone_id="NIA Mouse 7.5-dpc Whole Embryo cDNA Library
(Long)"
/note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
extracted from a pool of four embryos at 7.5-days
postcoitum. Double-stranded cDNAs were synthesized with an
Oligo(dT) primer (Invitrogen:
5'-GACTAGTCTAGATCGAGCGGCCGCCCTTTT-3') from
7 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker LR-SalI, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pSPORT1 plasmid vector.
The DH10B E. coli host was transformed with the ligation
mixture by the standard chemical method. The average
insert size is about 2.2 kb. The library was constructed
by Yulan Piao (NIA)."

BASE COUNT
ORIGIN

Query Match 69.5%; Score 204.4; DB 14; Length 503;
Best Local Similarity 81.0%; Pred. No. 1.6e-46;
Matches 238; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

1 ACATATGTGGCGCTCCATGTTGGACTAATTAGGAAGCTGCAGTTTATTGATCCT 60
|||||
Db 192 ACATTAATAGCAACCTTAATGTTGGAGTCTTAGAAGCTGTGATTTATTGATCCT 251
61 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAACATCTGTGTGATGATATAC 120
|||||
Db 252 CAAGAAGGTGGGAAGAAATTAGCATGTAGCTATCAAAAAGCCGTCCGGCAGCAGATAC 311
121 AATCAGTTTCAATTAAGAGATTGCAAGCTTACTTACAGACCGGAGAGACCCCACTGT 180
312 AATCAGTTTCAATTAAGAGATTGCAAGCTTACTTACAGACCGGAGAGACCCCACTGT 371
181 GAATTAAGTGTGACGCGGACCAAAATTTGCAATTTGATCTTTGGATCTTTTG 240
Db 372 GAATCTGTTTGACTGGGGACCAAGAACTGCAATTTGGCGACTTTGGATCTTACTG 431
241 ATCCAAATGAATTTTGTCTCTGCGAGTCTTTTGTCTCCAGATGCTTTCC 294

Db 432 GTCCAGATGAGCTGTTGGCCCCCGCACCTCTCTGCTGGCGATGCGTTCC 485

RESULT 8
B0552228
LOCUS
DEFINITION
H4014C09-5 NIA Mouse 7.4K cDNA clone set Mus musculus cDNA clone
H4014C09 5', mRNA sequence.
B0552228
ACCESSION
B0552228.1 GI:21453114
VERSION
B0552228.1
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (Bases 1 to 598)

REFERENCE
AUTHORS
VanBuren,V., Piao,Y., Dudekula,D.B., Qian,Y., Carter,M.G., Martin
P.R., Stagg,C.A., Baasey,U., Alba,K., Hamatani,T., Kargul,G.J.,
Luo,A.G., Kelso,J., Hide,W. and Ko,M.S.H.
Assembly, verification, and initial annotation of NIA 7.4K mouse
cDNA clone set
Genome Res. 12 (12), 1999-2003 (2002)
22354164
PUBMED
12466305

CONTACT: Yong Qian

Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://igsun.grc.nia.nih.gov/cDNA/NIA_7_4k.html for details.
Plate: H4014 row: C column: 05
Seq primer: -21M13 Reverse
High quality sequence stop: 598
POLYA=No.

FEATURES
SOURCE

Location/Qualifiers
1. 598
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="niaEST:H4014C09-5"
/db_xref="taxon:10090"
/clone="H4014C09"
/sex="mixed"
/dev_stage="mixed"
/lab_host="DH10B"
/clone_id="NIA Mouse 7.4K cDNA Clone Set"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
clone is among a rearranged set of 7,407 clones from more
than 20 cDNA libraries."

BASE COUNT
ORIGIN

Query Match 69.5%; Score 204.4; DB 13; Length 598;
Best Local Similarity 81.0%; Pred. No. 1.8e-46;
Matches 238; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

1 ACATATGTGGCGCTCCATGTTGGACTAATTAGGAAGCTGCAGTTTATTGATCCT 60
|||||
Db 176 ACATTAATAGCAACCTTAATGTTGGAGTCTTAGAAGCTGTGATTTATTGATCCT 235
61 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAACATCTGTGTGATGATATAC 120
|||||
Db 236 CAAGAAGGTGGGAAGAAATTAGCATGTAGCTATCAAAAAGCCGTCCGGCAGCAGATAC 295
121 AATCAGTTTCAATTAAGAGATTGCAAGCTTACTTACAGACCGGAGAGACCCCACTGT 180
296 AATCAGTTTCAATTAAGAGATTGCAAGCTTACTTACAGACCGGAGAGACCCCACTGT 355
181 GAATTAAGTGTGACGCGGACCAAAATTTGCAATTTGATCTTTGGATCTTTTG 240
Db 356 GAATCTGTTTGACTGGGGACCAAGAACTGCAATTTGGCGACTTTGGATCTTACTG 415

QY 241 ATCCAAATGAATTTTGTCTCTGCGAGCTTTTGTCTCCGAGATGCTGTC 294
 DB 416 GTTCAGATTGAGCTGTTTGCCCCCGCCACTCTCTCTGCGAGATGCTGTC 469
 RESULT 9
 LOCUS BB660378
 DEFINITION BB660378 RIKEN full-length enriched, 13 days embryo lung Mus
 BB660378 musculus cDNA clone D430042L21 5', mRNA sequence.
 ACCESION BB660378
 VERSION BB660378.1 GI:16494199
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 610)
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
 Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Sasaki,
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
 D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
 Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
 TITLE Unpublished
 JOURNAL Contact: Yoshihide Hayashizaki
 COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura,
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multichannel sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
 Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa,
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
 Hayashizaki, Y.
 Computational Analysis of Full-length cDNAs Compared with
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.
 e mouse tissues.
 location/Qualifiers
 1. 610
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="D430042L21"
 /issue_type="lung"
 /dev_stage="13 days embryo"
 /lab_host="DH10B"
 /clone_11b="RIKEN full-length enriched, 13 days embryo
 lung"

/note="Site 1: SalI; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5',
 GAGAGAGAGAGCGCCGACACTCGAGTGTGTTTTTTTTTTTNN 3'], cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. Second strand cDNA was prepared with the
 primer adapter of sequence [5',
 GAGAGAGAGATTCTCGAGTGAATTAATTAATTCCTCCCTCCCTCC 3']. cDNA
 was cleaved with BamHI and XhoI. Vector: a modified
 pBluescript KS(+) after bulk excision from Lambda FHC I."
 BASE COUNT. 149 a 179 c 164 g 118 t
 ORIGIN
 Query Match 69.5%; Score 204.4; DB 10; Length 610;
 Best Local Similarity 81.0%; Pred. No. 1.8e-46;
 Matches 238; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
 QY 1 ACATATGTCGCTGCTCAATGTGTGACTAATTAGAGAGCTGCAGATTATGATCCT 60
 DB 212 ACATCATATCGCAACCTTAATGTGGGATCCCTTAGAAGCTGTGAGATTATGATCCT 271
 QY 61 CAAGAAAGATGAGAGAGTGTAGCTATTAATAAACAATCTGTGATGATGATAC 120
 DB 272 CAAGAAAGGTGAGAGAAATTTAGCAATGCTATCAAAAAGCCCTCGGACACACATAC 331
 QY 121 AATCAGTTTCACATTAAGAGATTTGAAGCATTTACTTCAATCTGAGAAAAGTCCACTTCT 180
 DB 332 AATCAGTTTCATATAAGAGATTTGCAAGCTTACTTACAGACCGGAGAGGCCCACTGT 391
 QY 181 GAATTAATCTTTGACTGCGGACCAACAATTTGCAAGTTTGATCTTGATCTTTTG 240
 DB 392 GAATCTGCTTTGACTGCGGACCAACAATTTGCAAGTTTGATCTTGATCTTTTG 451
 QY 241 ATCCAAATGAATTTTGTCTCTGCGAGCTTTTGTCTCCGAGATGCTGTC 294
 DB 452 GTTCAGATTGAGCTGTTTGCCCCCGCCACTCTCTCTGCGAGATGCTGTC 505
 RESULT 10
 LOCUS BB613447
 DEFINITION BB613447 RIKEN full-length enriched, 10 day neonate skin Mus
 BB613447 musculus cDNA clone 4732482P03 5', mRNA sequence.
 ACCESION BB613447
 VERSION BB613447.1 GI:16454055
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 617)
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
 Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Sasaki,
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
 D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
 Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
 TITLE Unpublished
 JOURNAL Contact: Yoshihide Hayashizaki
 COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp,

to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCTCCAGTAATTAATTAATTCCTCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I"

BASE COUNT

180 a 196 c 172 g 128 t

Query Match 69.5%; Score 204.4; DB 10; Length 676;

Best Local Similarity 81.0%; Pred. No. 1,9e-46;

Matches 238; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

1 ACATATGTCGGCGCTCAATGTTGACATTAAGGAAGCTGCAGATTTATGATCT 60
140 ACATATGTCGGCGCTCAATGTTGACATTAAGGAAGCTGCAGATTTATGATCT 199
61 CAGAAGAGTGAAGAAGATTAGCTAGTATTAAAAACCATCTGTGTATGATAGATAC 120
200 CAGAAGAGTGAAGAAGATTAGCTAGTATTAAAAACCATCTGTGTATGATAGATAC 259
121 AATCAGTTTCACTAAGAGATTTGAGCTTCTTCAAACTGGAAGAGTCCACTTCT 180
260 AATCAGTTTCACTAAGAGATTTGAGCTTCTTCAAACTGGAAGAGTCCACTTCT 319
181 GAATTAAGTGTGTCAGGGGACACCAAAATTCAGAGTGTGTGATCTTTG 240
320 GAATTAAGTGTGTCAGGGGACACCAAAATTCAGAGTGTGTGATCTTTG 379
241 ATCCAAATGATTTTGTCTCTGCGAGTCTTTTCTCCAGATCTGTCTCC 294
380 GTCCAGATTCAGCTGTTTGGCCCCCGACCTCTCTGCGAGTCTGTCTCC 433

RESULT 14

AK020397 1161 bp mRNA linear HTC 05-DEC-2002

LOCUS Mus musculus adult male diencephalon cDNA, RIKEN full-length

DEFINITION enriched library, clone:933020903 product:interleukin-1

receptor-associated kinase 4 [Mus musculus], full insert sequence.

AK020397 GI:15919911

VERSION AK020397.2

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE Carninci, P. and Hayashizaki, Y.

JOURNAL High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

PUBMED 99279253

MEDLINE 10349636

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itom, H., Komuro, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

11042159

3

PUBMED

REFERENCE

AUTHORS

11078661

4

Kawai, J.,

Shinagawa, A.,

Shibata, K.,

Yoshino, M.,

Itom, M.,

Ishii, Y.,

Arakawa, T.,

Hara, A.,

Fukunishi, Y.,

Komuro, H.,

Adachi, J.,

Fukuda, S.,

Pukishi, Y.,

Puruno, M.,

Hanagaki, T.,

Hara, A.,

Hayatsu, N.,

Hiramoto, K.,

Hirooka, T.,

Horii, F.,

Imotani, K.,

Ishii, Y.,

Itom, H.,

Izawa, M.,

Kasukawa, T.,

Kato, H.,

Kawai, J.,

Kojima, Y.,

Komuro, H.,

Kouda, M.,

Koyama, S.,

Kuribara, C.,

Matsuyama, T.,

Myazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Shizuta, K., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

Direct Submission

Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail:genome-res@gscc.riken.go.jp,

URL: <http://genome.gsc.riken.go.jp>, Tel: 81-45-503-9222,

Fax: 81-45-503-9216)

On Oct 4, 2001 this sequence version replaced gi:12860970.

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

further details.

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues. First strand cDNA was primed with a primer

[5' GAGAGAGAGATCTCTCCAGTAATTAATTAATTCCTCCCCCCCC 3']. cDNA was

prepared by using triazole thermo-activated reverse transcriptase

and subsequently enriched for full-length by cap-trapper. cDNA went

through two rounds of normalization to Rot = 20.0 and subtraction

to Rot = 370.4. Second strand cDNA was prepared with the primer

adapter of sequence [5'

GAGAGAGAGATCTCTCCAGTAATTAATTAATTCCTCCCCCCCC 3']. cDNA was cleaved

with BamHI and XhoI. Vector: a modified pBluescript KS(+) after

bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI, 3'

end: BamHI. Host: DH10B

Location/Qualifiers

1. 1161

/organism="Mus musculus"

FEATURES

source

```

/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9330209D03"
/db_xref="MGI:1911393"
/db_xref="taxon:10090"
/clone="9330209D03"
/sex="male"
/tissue_type="diencephalon"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
183..776
/note="unnamed protein product; interleukin-1
receptor-associated kinase 4 [Mus musculus]
putative"
/codon_start=1
/protein_id="BAB32090.2"
/db_xref="GI:15919912"
/db_xref="MGI:1924812"
/translation="MNKPLPSTYIRNLNVLKLSDFIDPOGMRKLAVALKKGSG
DRYNOFHIRFEALLOTGKSPICELLPDGTTCVGDVLDLVLELPAATLLLP
DAVQIVKSLPPKAGNSGNTSLSGKODIRNVAKTRQIQLRATGLSPDRMSVES
SDTRFHSFSEHLEKSIINNFDDEOPASAGNRMEGPF"
BASE COUNT      317 a      296 c      308 g      240 t
ORIGIN
Query Match      69.5%; Score 204.4; DB 11; Length 1161;
Best Local Similarity 81.0%; Pred. No. 2.4e-46;
Matches 238; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1 ACATATGTGGCGCTCCATAGTTGACATTAATGGAAGCTGCAGATTTTATGATCT 60
    |||||
DB 207 ACATACATACGCAACCTTAATGTGGGATCCTTAGAAGCTGTCGATTTATGATCT 266
    |||||

QY 61 CAGAAGATGGAAGAGTAGCTGATTAATAAACACATGCTGATGATGATGATGATGAT 120
    |||||
DB 267 CAGAAGAGGGAAGAAATGACGATGATCAAAAAGCGTCCGACGACAGATGATGAT 326
    |||||

QY 121 AATCAGTTTCATTAAGAGATTTGAAGCATTAATCAAACTGGAAAAAGTCCACTCT 180
    |||||
DB 327 AATCAGTTTCATTAAGAGATTTGAAGCATTAATCAAACTGGAAAAAGTCCACTCT 386
    |||||

QY 181 GAATTAAGTGTGACCTGGGACCAACAATGACAGCTGATGATGATGATGATGATGAT 240
    |||||
DB 387 GAATGCTGTTGACCTGGGACCAACAAGTGAAGCTGATGATGATGATGATGATGAT 446
    |||||

QY 241 ATCCAAAATGAATTTTGTCTCTCGGACAGCTTTTGTCTCCCAATGCTGTTCC 294
    |||||
DB 447 GTCCAGATTAAGCTGTTGCTCCCGCACCTCTCTGCGGATGCGCTTCC 500
    |||||

RESULT 15
LOCUS      AK028837      2481 bp      mRNA      linear      HTC 05-DEC-2002
DEFINITION Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched
library, clone:473246109 product:interleukin-1 receptor-associated
kinase 4 [Mus musculus], full insert sequence.
ACCESSION  AK028837
VERSION     AK028837.1
KEYWORDS    HTc; CAP trapper.
SOURCE      Mus musculus
ORGANISM    Mus musculus (house mouse)
REFERENCE   1
  AUTHORS   Carninci, P. and Hayashizaki, Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
  PUBMED    10349636
REFERENCE   2
  AUTHORS   Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to

```

```

JOURNAL     Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE     20499374
PUBMED      11042159
REFERENCE
AUTHORS
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M.,
  Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
  Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
  Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watahiki, M.,
  Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, D.,
  Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
  RIKEN integrated sequence analysis (RISA) system-384-format
  sequencing pipeline with 384 multicapillary sequencer
  Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL     20530913
MEDLINE     11076861
PUBMED
REFERENCE
AUTHORS
4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
  Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
  Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
  Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
  Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
  Fleischmann, M., Gaasterland, T., Glass, C., King, B., Kochiwa, H.,
  Kuenl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
  Quackenbush, J., Schriml, L.M., Staudli, F., Suzuki, R., Tomita, M.,
  Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
  Balderrelli, R., Barsh, G., Blake, J., Brownstein, M.J., Bull, C.,
  Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C.,
  Fletcher, C., Fujita, M., Gariboldi, M., Guelinich, S., Hill, D.,
  Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
  Marchionni, L., Maehama, J., Mazzarelli, J., Mombarts, P., Nordone, P.,
  Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
  Sato, K., Schombach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
  Toyo-oka, K., Wang, K.H., Weltz, C., Whitaker, C., Wilming, L.,
  Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.
  and Hayashizaki, Y.
  Functional annotation of a full-length mouse cDNA collection
  Nature 409 (6821), 685-690 (2001)
JOURNAL     21085660
MEDLINE
PUBMED
REFERENCE
AUTHORS
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
  Group Phase I & II Team.
  Analysis of the mouse transcriptome based on functional annotation
  of 60,770 full-length cDNAs
  Nature 420, 563-573 (2002)
JOURNAL     6
MEDLINE
PUBMED
REFERENCE
AUTHORS
6 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
  Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
  Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
  Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
  Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
  Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murate, M.,
  Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
  Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
  Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
  Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
  Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
  Muramatsu, M. and Hayashizaki, Y.
  Direct Submision
  Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of
  Physical and Chemical Research (RIKEN), Laboratory for Genome
  Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
  RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
  Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
  URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
  Fax: 81-45-503-9216)
  CDNA library was prepared and sequenced in Mouse Genome
  Encyclopedia Project of Genome Exploration Research Group in Riken
  Genomic Sciences Center and Genome Science Laboratory in RIKEN.
  Division of Experimental Animal Research in Riken contributed to
  prepare mouse tissues.

```

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

Source

1. 2481

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:4732460109"

/db_xref="taxon:10090"

/clone="4732460109"

/tissue_type="skin"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="10 days neonate"

116..1495

/note="unnamed protein product; interleukin-1

receptor-associated kinase 4 [Mus musculus]

putative"

/codon_start=1

/protein_id="BAC26146.1"

/db_xref="GI:26324784"

/translation="MNKPLTPSTYIRLVNGLIRKLSDFIPDQGWKKLVAIKKPSG

DDRYNOFHIRFELALDTGKSPTECELLFDWGTNCTVGLVDLVLQIEFAPATLLP

DAVQIVKSLPPEAAVTAQTHQPCQKRTSVMPMKLEHSCPEPOSSSPDNRSVES

SDTRFHSFSEHLKSIITNFDQPSAGNMGSGGVVYKGVNNTVAIVKRLGAM

VEISTEELKQDFQDEIKVMAQCOHENLVELGFSDDNCLVYAVMPSGLDRLSC

LDGTPPLSMWTRCQVAGTANGIRFLHFNHHRDILKSANILDKDPTAKISDEGLAR

ASARIQTVMSRIVGTGTAAVPAEALRGELTPKSDIYSPGVVLELITGLAAVDENR

PQILLDIKERIEDEBETIEDYTBKMSDADPASVEAMYSASQCLHEKKRRPDIAKV

QQLIQEMSA"

BASE COUNT 650 a 592 c 634 g 605 t

ORIGIN

Query Match 69.5%; Score 204.4; DB 11; Length 2481;

Best Local Similarity 81.0%; Pred. No. 3.4e-46;

Matches 238; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

```
QY 1 ACATATGCGCTGCTCAATGTTGACATAATTAGAAAGCTGCAGATTTTATTTGATCCT 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 140 ACATACATAGCAACCTTAATGTGGGATCCTTAGAAAGCTGTGCGATTTTATTTGATCCT 199
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 CAAGAGGATGGAAGAGTGTAGCTATTAATAAAACCATCTGATGATGATAC 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 200 CAAGAAAGGTGGAAGAAATTAGCAGTAGCTATCAAAAAGCCGTCGGCAGACAGATAC 259
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 AATCAGTTTCAATAAGAGATTGAAGCACTTACTGAAACTGGAATAAAGTCCACTTCT 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 260 AATCAGTTTCAATAAGAGATTGAAGCCTTACTTCAAGCCGGAAAGCCCACTCTGT 319
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 GAATTACTGTTTGACTGGGACACCAAAATTGACAGTTGATCTTGTGATCTTTTG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 320 GAATGCTGTTTGACTGGGACACCAAGATGACAGATTTGCGACCTTGTGATCTTACTG 379
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 ATCCAAAATGAATTTTGTCTCTGCGAGCTTTTGTCTCCAGATGCTGTTCC 294
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 380 GTCCAGATTGAGCTGTTTGGCCCGCACCTCTCTGCTGCGGATGCCGTTCC 433
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: January 18, 2004, 00:39:20

Job time : 1319 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 20:02:06 ; Search time 191 Seconds
(without alignments)
4155.156 Million cell updates/sec

Title: US-10-001-254-5
Perfect score: 294
Sequence: 1 acatatgtcgcgtccca.....tgcctccagatcgtctccc 294

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

_N_Geneseq_19JUN03.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	294	100.0	294	AAD40074	Human IRAK4 DD (de
2	294	100.0	833	AAA09319	Human cancer assoc
3	294	100.0	1383	AAA00079	Human IRAK4 gene #
4	294	100.0	1668	AA576805	DNA encoding novel
5	294	100.0	2817	AAD40085	Human IRAK4 gene #
6	292.4	99.5	1383	AAD10197	Human interleukin-
7	241.6	82.2	501	AA561608	Lung small cell ca
8	204.4	69.5	1542	AAD10198	Mouse interleukin-

9	150.4	51.2	405	23	AA576803	DNA encoding novel
10	137.6	46.8	211	24	AAD40084	Human IRAK4 short
11	137.6	46.8	415	23	AA576802	DNA encoding novel
12	137.6	46.8	2213	22	AAH13798	Human cDNA sequenc
13	104.8	35.6	408	24	AB056733	Human colon cancer
14	42.6	14.5	2378	23	ABV23177	Human prostate exp
15	42.6	14.5	2378	23	ABV29017	Human prostate exp
16	40.4	13.7	1879	20	AA589865	L. esculentum germ
17	40.4	13.7	2024	20	AA589866	L. esculentum germ
18	40.4	13.7	2024	21	AA38935	Tomato germacrone
19	39.6	13.5	628	22	AAH87656	Peppermint plant o
20	36	12.2	2000	24	AB216713	Arabidopsis thalia
21	35.2	12.0	14041	22	AAH48024	Internal control B
22	33.8	11.5	3744	21	AAH70149	Plasmodium falcipt
23	33.8	11.5	580073	18	AA589840	Mycoplasma genital
24	33.6	11.4	831	25	ABX08179	S. pneumoniae type
25	33.6	11.4	2349	21	AAH05514	Streptococcus pneu
26	33.6	11.4	2393	21	AAV52223	Streptococcus pneu
27	33.6	11.4	2162598	25	AB556454	Streptococcus pneu
28	33.2	11.3	665	22	AAI15957	Human breast cance
29	33.2	11.3	1273	24	AB215600	Arabidopsis thalia
30	33	11.2	3564	24	ABN66233	Streptococcus poly
31	33	11.2	4300	22	AAH42107	Nucleotide sequenc
32	33	11.2	640681	24	ABA92787	Buchnera sp. genom
33	32.8	11.2	477	21	AA27964	Human secreted pro
34	32.8	11.2	2862	24	AA199844	Bacillus KSM-P358
35	32.8	11.2	3813	24	AA199855	Bacillus KSM-P358
36	32.8	11.2	202001	24	AB552506	Human transporter
37	32.6	11.1	1728	20	AAH99301	C. elegans CED-6 p
38	32.6	11.1	1832	24	AB234930	Human gene express
39	32.6	11.1	2971	22	AAH54085	S. epidermidis gen
40	32.6	11.1	4029	10	AAH91323	DNA encoding rabbi
41	32.6	11.1	4029	19	AAH94064	Rabbit growth horm
42	32.4	11.0	3353	23	ABV17967	Human prostate exp
43	32.4	11.0	358	23	ABV17729	Human prostate exp
44	32.4	11.0	451	23	ABV47458	Human prostate exp
45	32.4	11.0	492	23	ABV47522	Human prostate exp

ALIGNMENTS

RESULT 1	
AAD40074	standard, cDNA; 294 BP.
XX	XX
AC	AAD40074;
XX	XX
DT	22-OCT-2002 (first entry)
XX	XX
DE	Human IRAK4 DD (death domain) cDNA.
XX	XX
KW	Human; death domain; DD; death effector domain; DED; Chlamydia infection;
KW	NN-AIC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
KW	inflammation; allergy; autoimmunity; allograft rejection; cell division;
KW	immune-based pathology; fibrosis; arthritis; arthritis; graft versus host disease;
KW	immunosuppressive; gene therapy; antisense therapy; gene; ss.
XX	XX
OS	Homo sapiens.
XX	XX
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..294
FT	/tag= a
FT	/product= "Human IRAK4 DD"
FT	/note= "No start and stop codon"
FT	/partial
PN	MO200240680-A2.
XX	XX
PD	23-MAY-2002.
XX	XX
PF	15-NOV-2001; 2001MO-US44844.
XX	XX

PR 17-NOV-2000; 2000US-0715893.
PR 29-JUN-2001; 2001US-301889P.

XX (BURN-) BURHAM INST.

PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roch W,
PI Steiner-Jewen F;

XX WPI; 2002-500222/53.

DR P-PSDB; AAE24854.

XX New polypeptide comprising a death domain or death effector domain,
PT useful for discovery of drugs that suppress infection, inflammation,
PT allergy, sepsis, autoimmunity, allograft rejection and other diseases

XX Claim 18; Page 173-174; 209pp; English.

XX The invention relates to an isolated polypeptide comprising a death
CC domain (DD), death effector domain (DED) or NB-ARC domain. The invention
CC is useful for identifying a binding agent, preferably a protein or a drug
CC that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
CC domain from DAB3, IRAK4, CTDD (Chlamydia trachomatis DD protein), DED4 or
CC NIDD (NGFR-interacting Death Domain), with a candidate binding agent and
CC detecting the association of the domain and the candidate binding agent,
CC by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
CC chemical crosslinking, nuclear magnetic resonance (NMR), mass
CC spectroscopy (MS) and FPA. The invention is useful for modulating the
CC level of a cell process such as cell proliferation, cell adhesion, cell
CC stress responses, responses to microbial infection and B cell
CC immunoglobulin class switching, in particular apoptosis within a cell.
CC Antibody specifically reactive with CTDD DD of C. trachomatis, C.
CC muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the
CC CTDD DD protein is useful for detecting a Chlamydia infection. The
CC invention is useful for modulating the activity of oncogenic proteins,
CC for treating a pathology caused by the oncogenic proteins and for
CC treating bacterial infections by modulating the activity of bacterial
CC proteins. The protein and antibody specific for it are useful for
CC discovery of drugs that suppress infection, inflammation, allergy,
CC sepsis, autoimmunity, allograft rejection and other diseases. The protein
CC is useful for treating immune-based pathologies, pathologies associated
CC with cell division, inflammatory diseases such as sepsis, fibrosis,
CC arthritis, graft versus host disease. The invention is used in antisense
CC therapy and gene therapy. The present sequence is human IRAK4 DD cDNA.

XX Sequence 294 BP; 64 A; 55 C; 62 G; 93 T; 0 other;

Query Match 100.0%; Score 294; DB 24; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-77;

Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACATATGTCGGCTGCTCAATGTTGACTTAATAGAAAGCTGCAATTTTATGATCCT 60
DB 1 ACATATGTCGGCTGCTCAATGTTGACTTAATAGAAAGCTGCAATTTTATGATCCT 60
QY 61 CAAGAAGATGGAAGAAGTTAGCTGTATTAAAAAACAATCTGCTGATGATAGATAC 120
DB 61 CAAGAAGATGGAAGAAGTTAGCTGTATTAAAAAACAATCTGCTGATGATAGATAC 120
QY 121 AATCAGTTTACATTAAGAGATTGGAAGATTAATCAAACTGAAAAAGTCCACTTCT 180
DB 121 AATCAGTTTACATTAAGAGATTGGAAGATTAATCAAACTGAAAAAGTCCACTTCT 180
QY 181 GAATTTACTGTTGACTGCGGACCAACAATTGACACAGTTGATCTTTGGAATCTTTTG 240
DB 181 GAATTTACTGTTGACTGCGGACCAACAATTGACACAGTTGATCTTTGGAATCTTTTG 240
QY 241 ATCCAAATGAATTTTGTCTGCTGAGATCTTTTGTCCAGATGCTGTTCC 294
DB 241 ATCCAAATGAATTTTGTCTGCTGAGATCTTTTGTCCAGATGCTGTTCC 294

RESULT 2

AAA09319
ID AAA09319 standard; DNA; 833 BP.

XX AAA09319;

XX 10-AUG-2000 (first entry)

XX Human cancer associated antigen precursor DNA, clone NY-REN-64.

XX renal cancer; cancer associated antigen precursor; diagnosis;

XX cytostatic; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 50..670

XX MO200020587-A2.

XX 13-APR-2000.

XX 04-OCT-1999; 99WO-US22873.

XX 05-OCT-1998; 98US-0166300.

XX 05-OCT-1998; 98US-0166350.

XX (LUDWIG) LUDWIG INST CANCER RES.

XX Obata Y, Gout I, Threlk O, Sahin U, Pfeundschn M, Scanlan MJ;

XX Stockert E, Chen Y, Old LJ, Jager E, Knuth A;

XX WPI; 2000-303774/26.

XX P-PSDB; AAY92347.

XX abnormal expression of human cancer associated antigens

XX Claim 57; Page 85; 121pp; English.

XX AAA09310-20 are novel genes isolated by SEREX screening from a renal
CC cancer cell line 1973/10.4. The genes encode cancer associated antigen
CC precursors. These gene products are useful in methods for preventing,
CC diagnosing and/or treating disorders, especially cancer, associated with
CC abnormal expression of human cancer associated antigens. The method
CC comprises contacting a sample from a subject with an agent that
CC specifically binds to the nucleic acid molecule or expression product
CC (or fragment) complexed with a human leukocyte antigen (HLA) molecule
CC and determining the interaction between the agent and the nucleic acid
CC molecule or the expression product as a determination of the disorder.

XX Sequence 833 BP; 273 A; 155 C; 179 G; 226 T; 0 other;

Query Match 100.0%; Score 294; DB 21; Length 833;
Best Local Similarity 100.0%; Pred. No. 2.2e-77;

Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACATATGTCGGCTGCTCAATGTTGACTTAATAGAAAGCTGCAATTTTATGATCCT 60
DB 74 ACATATGTCGGCTGCTCAATGTTGACTTAATAGAAAGCTGCAATTTTATGATCCT 133
QY 61 CAAGAAGATGGAAGAAGTTAGCTGTATTAAAAAACAATCTGCTGATGATAGATAC 120
DB 61 CAAGAAGATGGAAGAAGTTAGCTGTATTAAAAAACAATCTGCTGATGATAGATAC 120
QY 121 AATCAGTTTACATTAAGAGATTGGAAGATTAATCAAACTGAAAAAGTCCACTTCT 180
DB 121 AATCAGTTTACATTAAGAGATTGGAAGATTAATCAAACTGAAAAAGTCCACTTCT 180
QY 181 GAATTTACTGTTGACTGCGGACCAACAATTGACACAGTTGATCTTTGGAATCTTTTG 240
DB 181 GAATTTACTGTTGACTGCGGACCAACAATTGACACAGTTGATCTTTGGAATCTTTTG 240
QY 254 GAATTTACTGTTGACTGCGGACCAACAATTGACACAGTTGATCTTTGGAATCTTTTG 313

QY 241 ATCCAAATGAAATTTTGTCTCTCGAGTCTTTTGTCTCCAGATGCTGTTCCC 294
 DB 314 ATCCAAATGAAATTTTGTCTCTCGAGTCTTTTGTCTCTCCAGATGCTGTTCCC 367

RESULT 3
 ID AAD40079 standard; DNA; 1383 BP.
 XX AAD40079;
 XX AAD40079;
 DT 22-OCT-2002 (first entry)
 DE Human IRAK4 gene #1.
 XX
 KW Human; death domain; DD; death effector domain; DED; Chlamydia infection;
 KW NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
 KW inflammation; allergy; autoimmunity; allograft rejection; cell division;
 KW immune-based pathology; fibrosis; arthritis; graft versus host disease;
 KW immunosuppressive; gene therapy; antisense therapy; gene; de.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..1383
 FT /*tag= a
 FT /product= "Human IRAK4"
 PN WO200240680-A2.
 XX 23-MAY-2002.
 PD
 XX 15-NOV-2001; 2001WO-US44844.
 PF 17-NOV-2000; 2000US-0715893.
 PR 29-JUN-2001; 2001US-301889P.
 XX
 PA (BURN-) BURNNAM INST.
 PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W,
 PI Stenmer-Jensen F;
 XX WPI; 2002-500222/53.
 DR P-PSDB; AAB24859.
 XX
 PT New polypeptide comprising a death domain or death effector domain,
 PT useful for discovery of drugs that suppress infection, inflammation,
 PT allergy, sepsis, autoimmunity, allograft rejection and other diseases
 PT
 XX
 PS Claim 19; Page 180-182; 209pp; English.

The invention relates to an isolated polypeptide comprising a death domain (DD), death effector domain (DED) or NB-ARC domain. The invention is useful for identifying a binding agent, preferably a protein or a drug that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC domain from DAP3, IRAK4, CTDD (Chlamydia trachomatis DD protein), DED4 or NIDD (NGFR-interacting Death Domain), with a candidate binding agent and detecting the association of the domain and the candidate binding agent, by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or chemical crosslinking, nuclear magnetic resonance (NMR), mass spectroscopy (MS) and FPA. The invention is useful for modulating the level of a cell process such as cell proliferation, cell adhesion, cell stress responses, responses to microbial infection and B cell immunoglobulin class switching, in particular apoptosis within a cell. Antibody specifically reactive with CTDD DD of C. trachomatis, C. muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the CTDD DD protein is useful for detecting a Chlamydia infection. The invention is useful for modulating the activity of oncogenic proteins, for treating a pathology caused by the oncogenic proteins and for treating bacterial infections by modulating the activity of bacterial proteins. The protein and antibody specific for it are useful for discovery of drugs that suppress infection, inflammation, allergy,

CC sepsis, autoimmunity, allograft rejection and other diseases. The protein is useful for treating immune-based pathologies, pathologies associated with cell division, inflammatory diseases such as sepsis, fibrosis, arthritis, graft versus host disease. The invention is used in antisense therapy and gene therapy. The present sequence is human IRAK4 gene.

SQ Sequence 1383 BP; 463 A; 243 C; 283 G; 394 T; 0 other;

Query Match 100.0%; Score 294; DB 24; Length 1383;
 Best Local Similarity 100.0%; Pred. No. 2.6e-77;
 Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACATATGCGCTGCTCTCAATGTTGACTAATTAGAAAGCTGATGATTTATGATCCT 60
 DB 25 ACATATGCGCTGCTCTCAATGTTGACTAATTAGAAAGCTGATGATTTATGATCCT 84
 QY 61 CAAGAAGATGAAAGAAAGTTAGCTGTAGCTATTTAAAACCATCTGTGTGATGATGATAC 120
 DB 85 CAAGAAGATGAAAGAAAGTTAGCTGTAGCTATTTAAAACCATCTGTGTGATGATGATAC 144
 QY 121 AATCATGTTTCAATTAAGAGATTTGAAGCATTTACTTCAAACTGGAAGAAAGTCCCATCTCT 180
 DB 145 AATCATGTTTCAATTAAGAGATTTGAAGCATTTACTTCAAACTGGAAGAAAGTCCCATCTCT 204
 QY 181 GAATTACTGTTTGACTGGGACCAACAATTTGACAGTTGATGATCTTTG 240
 DB 205 GAATTACTGTTTGACTGGGACCAACAATTTGACAGTTGATGATCTTTG 264
 QY 241 ATCCAAATGAAATTTTGTCTCTCGAGTCTTTTGTCTCCAGATGCTGTTCCC 294
 DB 265 ATCCAAATGAAATTTTGTCTCTCGAGTCTTTTGTCTCCAGATGCTGTTCCC 318

RESULT 4
 AAS76805
 ID AAS76805 standard; cDNA; 1668 BP.
 XX
 AC AAS76805;
 DT 13-FEB-2002 (first entry)
 DE DNA encoding novel human diagnostic protein #12609.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSB-) HYSBQ INC.
 XX
 PI Dmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG12618.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PT
 XX
 PS Claim 1; SEQ ID No 12609; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. A564197-A594564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1668 BP; 571 A; 289 C; 332 G; 476 T; 0 other;
Query Match 100.0%; Score 294; DB 23; Length 1668;
Best Local Similarity 100.0%; Pred. No. 2.8e-77;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACATATGTCGCGCTCAATGTTGACCTAATTAGGAAGCTGTCAGATTATTGATCCT 60
DB 25 ACATATGTCGCGCTCAATGTTGACCTAATTAGGAAGCTGTCAGATTATTGATCCT 84
QY 61 CAGAAGAGTGAAGAAGATTAGCTGTAGCTATTAAAAACATCTGGTGTATGATAGTAC 120
DB 85 CAGAAGAGTGAAGAAGATTAGCTGTAGCTATTAAAAACATCTGGTGTATGATAGTAC 144
QY 121 AATCAGTTTACATAGGAAGATTGAGCACTTACTTAACCTGGAAAAATGCCACTTCT 180
DB 145 AATCAGTTTACATAGGAAGATTGAGCACTTACTTAACCTGGAAAAATGCCACTTCT 204
QY 181 GAATTAAGTTTACATAGGAAGATTGAGCACTTACTTAACTGGTGTATGATAGTAC 240
DB 205 GAATTAAGTTTACATAGGAAGATTGAGCACTTACTTAACTGGTGTATGATAGTAC 264
QY 241 ATCCAAAATGAATTTTGTCTCTCGAGAGCTTTTGTCTCCAGATGCTGTCC 294
DB 265 ATCCAAAATGAATTTTGTCTCTCGAGAGCTTTTGTCTCCAGATGCTGTCC 318
RESULT 5
AAD40085
ID AAD40085 standard; DNA; 2817 BP.
XX
AC AAD40085;
XX
DT 22-OCT-2002 (first entry)
XX
DE Human IRAK4 gene #2.
XX
KW Human; death domain; DD; death effector domain; DED; Chlamydia infection;
KW NB-ARC domain; apoptosis; oncogenic protein; Bacterial infection; sepsis;
KW inflammation; allergy; autoimmunity; allograft rejection; cell division;
KW immune-based pathology; fibrosis; arthritis; graft versus host disease;
KW immunosuppressive; gene therapy; antisense therapy; gene; ds.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT 50..1432
XX FT CDS /*tag= a
XX FT /product= "Human IRAK4"
XX PN MO200240680-A2.
XX PD 23-MAY-2002.

XX
XX 15-NOV-2001; 2001WO-US44844.
XX
XX 17-NOV-2000; 2000US-0715893.
XX 29-UTN-2001; 2001US-301889P.
XX
XX (BURN-) BURNHAM INST.
XX
XX Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
XX Steiner-Lieman F;
XX
XX WPI, 2002-500222/53.
XX P-PSDB; AAE24865.
XX
XX New polypeptide comprising a death domain or death effector domain,
XX useful for discovery of drugs that suppress infection, inflammation,
XX allergy, sepsis, autoimmunity, allograft rejection and other diseases
XX
XX Claim 19; Page 194-196; 209pp; English.
XX
XX The invention relates to an isolated polypeptide comprising a death
XX domain (DD), death effector domain (DED) or NB-ARC domain. The invention
XX is useful for identifying a binding agent, preferably a protein or a drug
XX that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
XX domain from DAP3, IRAK4, CTDD (Chlamydia trachomatis DD protein), DED4 or
XX NIDD (NGFR-interacting Death Domain), with a candidate binding agent and
XX detecting the association of the domain and the candidate binding agent,
XX by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
XX chemical crosslinking, nuclear magnetic resonance (NMR), mass
XX spectroscopy (MS) and FPA. The invention is useful for modulating the
XX level of a cell process such as cell proliferation, cell adhesion, cell
XX stress responses, responses to microbial infection and B cell
XX immunoglobulin class switching, in particular apoptosis within a cell.
XX Antibody specifically reactive with CTDD DD of C. trachomatis, C.
XX muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the
XX CTDD DD protein is useful for detecting a Chlamydia infection. The
XX invention is useful for modulating the activity of oncogenic proteins,
XX for treating a pathology caused by the activity of oncogenic proteins,
XX for treating bacterial infections by modulating the activity of bacterial
XX proteins. The protein and antibody specific for it are useful for
XX discovery of drugs that suppress infection, inflammation, allergy,
XX sepsis, autoimmunity, allograft rejection and other diseases. The protein
XX is useful for treating immune-based pathologies, pathologies associated
XX with cell division, inflammatory diseases such as sepsis, fibrosis,
XX arthritis, graft versus host disease. The invention is used in antisense
XX therapy and gene therapy. The present sequence is human IRAK4 gene.
XX
SQ Sequence 2817 BP; 912 A; 547 C; 586 G; 772 T; 0 other;
Query Match 100.0%; Score 294; DB 24; Length 2817;
Best Local Similarity 100.0%; Pred. No. 3.3e-77;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACATATGTCGCGCTCAATGTTGACCTAATTAGGAAGCTGTCAGATTATTGATCCT 60
DB 74 ACATATGTCGCGCTCAATGTTGACCTAATTAGGAAGCTGTCAGATTATTGATCCT 133
QY 61 CAGAAGAGTGAAGAAGATTAGCTGTAGCTATTAAAAACATCTGGTGTATGATAGTAC 120
DB 134 CAGAAGAGTGAAGAAGATTAGCTGTAGCTATTAAAAACATCTGGTGTATGATAGTAC 193
QY 121 AATCAGTTTACATAGGAAGATTGAGCACTTACTTAACTGGTGTATGATAGTAC 180
DB 194 AATCAGTTTACATAGGAAGATTGAGCACTTACTTAACTGGTGTATGATAGTAC 253
QY 181 GAATTAAGTTTACATAGGAAGATTGAGCACTTACTTAACTGGTGTATGATAGTAC 240
DB 254 GAATTAAGTTTACATAGGAAGATTGAGCACTTACTTAACTGGTGTATGATAGTAC 313
QY 241 ATCCAAAATGAATTTTGTCTCTCGAGAGCTTTTGTCTCCAGATGCTGTCC 294
DB 314 ATCCAAAATGAATTTTGTCTCTCGAGAGCTTTTGTCTCCAGATGCTGTCC 367

RESULT 6
AADI0197
ID AADI0197 standard; cDNA; 1383 BP.
XX
AC AADI0197;
XX
DT 24-SEP-2001 (first entry)
XX
DE Human interleukin-1 (IL-1) receptor-associated kinase (IRAK)-4 cDNA.
XX
KM Human; interleukin-1 receptor associated kinase-4; IRAK-4; cytosolic;
KM IL; antibacterial; antiinflammatory; ophthalmological; vasotropic; OPD;
KM chronic obstructive pulmonary disease; neuroprotective; chronic cough;
KM adult respiratory disease syndrome; pulmonary fibrosis; asthma; ARDS;
KM interstitial lung disease; allergic rhinitis; transplant rejection;
KM autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;
KM multiple sclerosis; diabetes; cancer; solid tumour; lymphoma; stroke;
KM cardiovascular disease; atherosclerosis; neurodegenerative disease;
KM sepsis; osteoarthritis; rash; osteoporosis; psoriasis; dermatitis;
KM inflammatory bowel disease; gout; Crohn's disease; ulcerative colitis;
KM Behcet's syndrome; ankylosing spondylitis; IL-1 receptor/Toll receptor;
KM sarcoidosis; transgenic animal; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1383
FT /*tag= a
FT /*product= "Human IRAK-4"
XX
XX WO200151641-A1.
XX
PN 19-JUL-2001.
XX
PF 12-JAN-2001; 2001WO-US01171.
XX
PR 13-JAN-2000; 2000US-0176395.
XX
PA (TULA-) TULARIK INC.
XX
PI Wesche H, Li S;
XX
DR WPI: 2001-451860/48.
XX
DR P-PSDB; AAB05398.
XX
PT Novel human interleukin-1 receptor associated kinase polypeptide,
PT useful for identifying modulators of the polypeptide for treating gout,
PT asthma, allergic rhinitis, multiple sclerosis and skin cancer -
XX
XX Claim 7; Fig 2; 89pp; English.
XX
XX The present sequence is a cDNA encoding human interleukin (IL)-1 receptor
XX associated kinase (IRAK)-4. IRAK associate with activated IL-1, IL-18
XX and other receptors and act to transduce signals originating from the
XX activated receptors, ultimately leading to a variety of downstream
XX effects such as nuclear factor (NF)-kappaB activation. The IRAK-4
XX inhibitors are useful for treating inflammatory diseases such as
XX pulmonary diseases and diseases of the airway (e.g., adult respiratory
XX disease syndrome (ARDS), chronic obstructive pulmonary disease (OPD),
XX pulmonary fibrosis, interstitial lung disease, asthma, chronic cough,
XX or allergic rhinitis), transplant rejection, autoimmune diseases (e.g.,
XX rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis
XX or diabetes), cancer (e.g., solid tumours, skin cancer or lymphoma),
XX cardiovascular diseases (e.g., stroke and atherosclerosis), diseases
XX of the central nervous system (e.g., neurodegenerative disease), CD14
XX mediated sepsis, non-CD14 mediated sepsis, osteoarthritis, osteoporosis,
XX psoriasis, diseases of the skin (e.g., rash, contact dermatitis, atopic
XX dermatitis), inflammatory bowel disease (e.g., Crohn's disease and
XX ulcerative colitis), Behcet's syndrome, ankylosing spondylitis, gout,
XX sarcoidosis and ophthalmic diseases and conditions. The inhibitors of
XX IRAK-4 activity or expression are used to inhibit signal transduction

CC resulting from the activation of an interleukin-1 receptor (IL-1R)/Toll
CC receptor in a cell. They also inhibit the activation of a transcription
CC factor that activates NF-kappaB in the cell. IRAK-4 is used to create a
CC nonhuman transgenic animal which is useful for testing the function of
CC IRAK-4 in vivo, to generate models for the study of inflammatory
CC disorders and conditions and for the development of potential treatments
CC for IRAK-4-related inflammatory diseases and conditions. IRAK-4 sequences
CC are also used in gene therapy and in antisense therapy.
XX
SQ Sequence 1383 BP; 463 A; 244 C; 283 T; 0 other;
XX
Query Match 99.5%; Score 292.4; DB 22; Length 1383;
Best Local Similarity 99.7%; Pred. No. 7.8e-77;
Matches 293; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 ACATATGTCGCTGCTCAATGTTGACTATTAATGAGAGCTGCAGATTTTATGATCCT 60
XX
DB 25 ACATATGTCGCTGCTCAATGTTGACTATTAATGAGAGCTGCAGATTTTATGATCCT 84
XX
QY 61 CAAAGAGATGAGAAAGATTAGCTGTAGCTATTAATAAACCATCTGTGATGATGATAC 120
XX
DB 85 CAAAGAGATGAGAAAGATTAGCTGTAGCTATTAATAAACCATCTGTGATGATGATAC 144
XX
QY 121 AATCAGTTTCACATTAAGAGATTGAAGCATTTACTTCAAACTGAAAAAGTCCACCTCT 180
XX
DB 145 AATCAGTTTCACATTAAGAGATTGAAGCATTTACTTCAAACTGAAAAAGTCCACCTCT 204
XX
QY 181 GAATTACTGTTTACTGCTGGGACACCAAAATTTGACAGTGTGATCTTTGATCTTTTG 240
XX
DB 205 GAATTACTGTTTACTGCTGGGACACCAAAATTTGACAGTGTGATCTTTGATCTTTTG 264
XX
QY 241 ATCCAAATTAATTTTGTCTGCTGGAGCTTTTGTCTCCCAATGCTGTCTCC 294
XX
DB 265 ATCCAAATTAATTTTGTCTGCTGGAGCTTTTGTCTCCCAATGCTGTCTCC 318
XX
RESULT 7
AAS61608
ID AAS61608 standard; cDNA; 501 BP.
XX
XX AAS61608;
XX
AC AAS61608;
XX
DT 29-JAN-2002 (first entry)
XX
XX
DE Lung small cell carcinoma antigen, cDNA #149.
XX
KM Human; cytostatic; antitumour; lung small cell cancer antigen;
KM tumour; lung cancer; ss.
XX
OS Homo sapiens.
XX
XX WO200177168-A2.
XX
PN 18-OCT-2001.
XX
PD 11-APR-2001; 2001WO-US11859.
XX
PF 11-APR-2000; 2000US-196780P.
XX
PR 21-JUN-2000; 2000US-213361P.
XX
PR 01-SEP-2000; 2000US-229763P.
XX
PR 05-SEP-2000; 2000US-230629P.
XX
PR 14-SEP-2000; 2000US-232565P.
XX
PR 19-DEC-2000; 2000US-257037P.
XX
PR 08-JAN-2001; 2001US-260796P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Lodes MJ, Wang T, Mohamath R, Indrias CV;
XX
DR WPI: 2002-010896/01.
XX
PT Lung tumour polynucleotide and polypeptides useful in therapy and
PT diagnosis of cancer especially lung cancer -

QY 181 GAATTACTGTTTGAAGTGGGACCAAAATTGACAGTTGGTGTATCTTTGGATCTTTTG 240
DB 367 GAATCGCTGTTGACTGGGACCAACGAACTGACAGTTGGCGACTTTGGATCTTACTG 426
QY 241 ATCCAAATGAAATTTTGTCTCTCCGAGTCTTTTGTCTCCCAAGTGTCTTCCC 294
DB 427 GTCCAGATTGAGCTTTTGGCCCCCGCCACTCTCTGCTGCGGATGCGTTC 480

RESULT 9
ID AAS76803 standard; cDNA; 405 BP.
AC AAS76803;
XX
XX 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #12607.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
OS
XX MO200175067-A2.
PN
XX 11-OCT-2001.
PD
PF 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HXSE-) HXSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI; 2001-639362/73.
DR P-PSDB; ABG12616.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX Claim 1; SEQ ID No 12607; 103bp; English.
PS
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridization probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations in
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 405 BP; 93 A; 87 C; 83 G; 142 T; 0 other;

Query Match 51.2%; Score 150.4; DB 23; Length 405;

Best Local Similarity 99.3%; Pred. No. 7.9e-35;
Matches 151; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 134 TAAGAGATTTGAAGATTTACTTCAAACTGAAAGTCCCATCTTCTGATTTCTGTTG 193
DB 115 TAAGAGATTTGAAGATTTACTTCAAACTGAAAGTCCCATCTTCTGATTTCTGTTG 174
QY 194 ACTGGGACACCAAAATTGACAGTTGTGATCTTGTGATCTTTGATCCAAATGAAT 253
DB 175 ACTGGGACACCAAAATTGACAGTTGTGATCTTGTGATCTTTGATCCAAATGAAT 234
QY 254 TTTTGTCTCTGCGAGTCTTTTGTCTCCAGAT 285
DB 235 TTTTGTCTCTGCGAGTCTTTTGTCTCCAGAT 266

RESULT 10
ID AAD40084 standard; DNA; 211 BP.
AC AAD40084;
XX
XX 22-OCT-2002 (first entry)
DE
XX Human IRAK4 short gene.
XX
XX Human; death domain; DD; death effector domain; DED; Chlamydia infection;
KM NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
KM inflammation; allergy; autoimmunity; allograft rejection; cell division;
KM immune-based pathology; fibrosis; arthritis; graft versus host disease;
KM immunosuppressive; gene therapy; antisense therapy; gene; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1..180
FT CDS /*tag= a
FT /product= "Human IRAK4 short protein"
XX
XX NO200240680-A2.
PN
XX 23-MAY-2002.
PD
XX 15-NOV-2001; 2001WO-US44844.
PF
XX 17-NOV-2000; 2000US-0715893.
PR 29-JUN-2001; 2001US-301889P.
XX
XX (BURN-) BURHAM INST.
PA
XX
PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W,
PI Semner-Liwen F;
XX
XX WPI; 2002-500222/53.
DR P-PSDB; AAE24864.
XX
XX New polypeptide comprising a death domain or death effector domain,
PT useful for discovery of drugs that suppress infection, inflammation,
PT allergy, sepsis, autoimmunity, allograft rejection and other diseases
PT
XX
PS Claim 19; Page 195; 209pp; English.
XX
XX The invention relates to an isolated polypeptide comprising a death
CC domain (DD), death effector domain (DED) or NB-ARC domain. The invention
CC is useful for identifying a binding agent, preferably a protein or a drug
CC that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
CC domain from DAP3, IRAK4, CRPD (Chlamydia trachomatis DD protein), DED4 or
CC NIDD (NGFR-interacting Death Domain), with a candidate binding agent and
CC detecting the association of the domain and the candidate binding agent,
CC by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
CC chemical crosslinking, nuclear magnetic resonance (NMR), mass
CC spectroscopy (MS) and FPA. The invention is useful for modulating the

CC level of a cell process such as cell proliferation, cell adhesion, cell
CC stresses responses, responses to microbial infection and B cell
CC immunoglobulin class switching, in particular apoptosis within a cell.
CC Antibody specifically reactive with CTDD DD of C. trachomatis, C.
CC muridarum, C. pneumoniae, and C. pitteci or a nucleic acid encoding the
CC CTDD DD protein is useful for detecting a Chlamydia infection. The
CC invention is useful for modulating the activity of oncogenic proteins,
CC for treating a pathology caused by the oncogenic proteins and for
CC treating bacterial infections by modulating the activity of bacterial
CC proteins. The protein and antibody specific for it are useful for
CC discovery of drugs that suppress infection, inflammation, allergy,
CC sepsis, autoimmunity, allograft rejection and other diseases. The protein
CC is useful for treating immune-based pathologies, pathologies associated
CC with cell division, inflammatory diseases such as sepsis, fibrosis,
CC arthritis, graft versus host disease. The invention is used in antisense
CC therapy and gene therapy. The present sequence is human IRAK4 short gene.
XX
SQ Sequence 211 BP; 77 A; 41 C; 36 G; 57 T; 0 other;

Query Match 46.8%; Score 137.6; DB 24; Length 211;
Best Local Similarity 97.2%; Pred. No. 4.1e-31;
Matches 140; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACATATGTCGCTGCTCAATGTTGACATAATAGGAAGCTGCAAGTTTATTGATCCT 60
DB 25 ACATATGTCGCTGCTCAATGTTGACATAATAGGAAGCTGCAAGTTTATTGATCCT 84
QY 61 CAGAAGAGATGGAAGAAGTTAGCTAGCTATTAATAAACCATCTGCTGATGATGATAC 120
DB 85 CAGAAGAGATGGAAGAAGTTAGCTAGCTATTAATAAACCATCTGCTGATGATGATAC 144
QY 121 AATCAGTTTCACATTAAGAGATT 144
DB 145 AATCAGTTTCACATTAAGATCTGT 168

RESULT 11

AA576802 standard; cDNA; 415 BP.

AA576802;

13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #12606.

Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Dmanac RT, Liu C, Tang YT;

WPI, 2001-639362/73.

P-PSDB; ABG12615.

New isolated polynucleotide and encoded polypeptides, useful in
PT diagnosis, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

Claim 1; SEQ ID No 12606; 1033p; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AA564197-AA594564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 415 BP; 144 A; 83 C; 78 G; 110 T; 0 other;

Query Match 46.8%; Score 137.6; DB 23; Length 415;
Best Local Similarity 97.2%; Pred. No. 5.1e-31;
Matches 140; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACATATGTCGCTGCTCAATGTTGACATAATAGGAAGCTGCAAGTTTATTGATCCT 60
DB 60 ACATATGTCGCTGCTCAATGTTGACATAATAGGAAGCTGCAAGTTTATTGATCCT 119
QY 61 CAGAAGAGATGGAAGAAGTTAGCTAGCTATTAATAAACCATCTGCTGATGATGATAC 120
DB 120 CAGAAGAGATGGAAGAAGTTAGCTAGCTATTAATAAACCATCTGCTGATGATGATAC 179
QY 121 AATCAGTTTCACATTAAGAGATT 144
DB 180 AATCAGTTTCACATTAAGATCTGT 203

RESULT 12

AAH13798 standard; cDNA; 2213 BP.

AAH13798;

26-JUN-2001 (first entry)

Human cDNA sequence SEQ ID NO:10742.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELT-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX

DR WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 10742; 2537bp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification, where a primer set
CC comprises: (a) an oligo-dt primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH1628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 2213 BP; 642 A; 488 C; 516 G; 567 T; 0 other;
XX
Query Match 46.8%; Score 137.6; DB 22; Length 2213;
Best Local Similarity 97.2%; Pred. No. 8.9e-31;
Matches 140; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
QY 1 ACATATGTCGCTGCTCAATGTTGACTAATTAAGAAAGCTGATTTATGATCCT 60
DB |||||||||||||||||||||||||||||||||||||||||||||||||||
DB 775 ACATATGTCGCTGCTCAATGTTGACTAATTAAGAAAGCTGATTTATGATCCT 834
QY 61 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAAACCATCTGATGATAGATAC 120
DB |||||||||||||||||||||||||||||||||||||||||||||||||||
DB 835 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAAACCATCTGATGATAGATAC 894
QY 121 AATCAGTTTCACATTAAGAGATTT 144
DB |||||||||||||||||||||||||||||||||||||||||||||||||||
DB 895 AATCAGTTTCACATTAAGAGATTT 918
XX
RESULT 13
AB056733
ID AB056733 standard; cDNA; 408 BP.
XX
AC AB056733;
XX
XX 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:428.
XX
XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KM genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200229086-A2.
XX
XX 11-APR-2002.
XX
PD 02-OCT-2001; 2001WO-US30732.
XX
PF 02-OCT-2000; 2000US-237271P.
XX
PR

XX
PA (FARB) BAYER CORP.
XX
XX Bugges C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiaalingam A, Lewis ME;
XX
XX WPI; 2002-426115/45.
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT or tissue useful for determining the presence of colon cancer in a cell
PT or tissue type, and in antisense therapy -
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC AB056306 to AB060787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the AB060776 to AB060787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC microarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensic, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists.
XX
SQ Sequence 408 BP; 114 A; 92 C; 81 G; 61 T; 60 other;
XX
Query Match 35.6%; Score 104.8; DB 24; Length 408;
Best Local Similarity 69.5%; Pred. No. 2.8e-21;
Matches 207; Conservative 0; Mismatches 81; Indels 10; Gaps 7;
XX
QY 1 ACATATGTCGCTGCTCAATGTTGACTAATTAAGAAAGCTGATTTATGATCCT 60
DB |||||||||||||||||||||||||||||||||||||||||||||||||||
DB 39 ACATGTCGCTGCTGCTCAATGTTGACTAATTAAGAAAGCTGATTTATGATCCT 98
QY 61 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAAACCATCTGATGATAGATAC 119
DB |||||||||||||||||||||||||||||||||||||||||||||||||||
DB 99 NNANNAAGANGG-ANNAAGGAGCTGAGCTATTAATAAACCAACGCGGATGATNNATA 157
QY 120 CAATCAGTTTCACATA-AGGAGATTGAGCAATTACTTCAA--CTGAAAAAAGTCCCA 175
DB |||||||||||||||||||||||||||||||||||||||||||||||||||
DB 158 CCATCAGTTTCACANANNGAGAGATTGAGCAATTACTTCAA--CTGAAAAAAGTCCCA 217
QY 176 CTTCTGAAT-TACTGTTTGA-CTGGGGGACACCAAAATTGCAAGTT-AGTATCTTGTGGA 233
DB |||||||||||||||||||||||||||||||||||||||||||||||||||
DB 218 CTTGGAAGAAAGTCTGTTGCTGGGGACCAAAATTGCAAGTATGAGGATCTTGTGGA 277
QY 234 TCTTTTGATCCAAAATGAATTTTGTG--CTCTGCGAGTCTTTTGTGCTCCAGATGCTG 289
DB |||||||||||||||||||||||||||||||||||||||||||||||||||
DB 278 TNCNCGAANCAACACAGATTTTNTGCTGCTGCTGCAANGCGGCGGCCAGACGCG 335
XX
RESULT 14
ABV23177
ID ABV23177 standard; cDNA; 2378 BP.
XX
AC ABV23177;
XX
XX 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 23168.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX

PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 DR WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 PS Claim 1; Page 4160-4161; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 2378 BP; 769 A; 396 C; 487 G; 706 T; 20 other;
 SQ
 Query Match 14.5%; Score 42.6; DB 23; Length 2378;
 Best Local Similarity 49.8%; Pred. No. 0.015;
 Matches 108; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
 QY 59 CTCAGAGAGATGAGAGAAAGTTAGCTGATTAATAAAACCATCTGGATGATAGAT 118
 DB 467 CACAGAAATTTAGAAAGCTGCTTTCTAAGAACTATCCACCAATAGATCAAGTTATAC 526
 QY 119 ACAATCAGTTTCACATAGAGAGATTTGAAGCATTAACCTGAAAGAAAGTCCACTT 178
 DB 527 AGAAACACAGAGATTGTACAGAGATTGTGAATTTCTTGAAAGAAATGCAATTTGCACTT 586
 QY 179 CTGAATTACTGTTTGAAGCTGGGGACACAAATTCAGATTTGGATCTTGGATCTTT 238
 DB 587 TACAATTGAAAGCTGCATGGGCAATTAACAATATAGATCTGGAACCTTTTCTGCATACCA 646
 QY 239 TGATCCAAATGAATTTTTCCTCCTCGAGTCTTTT 275
 DB 647 AGGTAGTATGAAACTGGGGCTGTTCGATTTTAT 683
 RESULT 15
 ABV29017
 ID ABV29017 standard; cDNA; 2378 BP.
 XX AC
 XX ABV29017;
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 29008.

XX
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX pharmacogenomic marker; gene; ss.
 XX Homo sapiens.
 OS
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 DR WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 PS Claim 1; Page 6140; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 2378 BP; 769 A; 396 C; 487 G; 706 T; 20 other;
 SQ
 Query Match 14.5%; Score 42.6; DB 23; Length 2378;
 Best Local Similarity 49.8%; Pred. No. 0.015;
 Matches 108; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
 QY 59 CTCAGAGAGATGAGAGAAAGTTAGCTGATTAATAAAACCATCTGGATGATAGAT 118
 DB 467 CACAGAAATTTAGAAAGCTGCTTTCTAAGAACTATCCACCAATAGATCAAGTTATAC 526
 QY 119 ACAATCAGTTTCACATAGAGAGATTTGAAGCATTAACCTGAAAGAAAGTCCACTT 178
 DB 527 AGAAACACAGAGATTGTACAGAGATTGTGAATTTCTTGAAAGAAATGCAATTTGCACTT 586
 QY 179 CTGAATTACTGTTTGAAGCTGGGGACACAAATTCAGATTTGGATCTTGGATCTTT 238
 DB 587 TACAATTGAAAGCTGCATGGGCAATTAACAATATAGATCTGGAACCTTTTCTGCATACCA 646
 QY 239 TGATCCAAATGAATTTTTCCTCCTCGAGTCTTTT 275
 DB 647 AGGTAGTATGAAACTGGGGCTGTTCGATTTTAT 683
 Search completed: January 17, 2004, 23:50:44
 Job time : 195 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2004, 00:39:27 ; Search time 214 Seconds
(without alignments)
4842.432 Million cell updates/sec

Title: US-10-001-254-5
Perfect score: 294
Sequence: 1 acatattgctgcctcctcaaa.....tgctccagatgctgtcc 294

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	294	100.0	294	US-10-001-254-5	Sequence 5, Appl1
2	294	100.0	1383	US-10-001-254-15	Sequence 15, Appl1
3	294	100.0	2817	US-09-966-451-3	Sequence 3, Appl1
4	294	100.0	2817	US-10-001-254-27	Sequence 27, Appl1
5	292.4	99.5	1383	US-09-759-595-2	Sequence 2, Appl1
6	281.6	82.2	501	US-09-833-790-149	Sequence 149, App
7	204.4	69.5	1542	US-09-759-595-4	Sequence 4, Appl1
8	151.8	51.6	31000	US-09-966-451-10	Sequence 10, Appl1
9	137.6	46.8	211	US-10-001-254-25	Sequence 25, Appl1
10	40.4	13.7	2024	US-09-887-586A-51	Sequence 51, Appl1
11	40.4	13.7	2024	US-09-903-012-51	Sequence 51, Appl1
12	40.4	13.7	2024	US-09-900-797-51	Sequence 51, Appl1
13	36	12.2	2000	US-09-938-842A-518	Sequence 4518, Ap
14	33.8	11.5	22684	US-09-960-858-2	Sequence 2, Appl1
15	33.8	11.5	22684	US-09-960-870-2	Sequence 2, Appl1

C 16	33.8	11.5	58073	13	US-10-205-220-1	Sequence 1, Appl1
C 17	33.6	11.4	2349	11	US-09-769-787-287	Sequence 287, App
C 18	33.2	11.3	899	15	US-10-198-846-7068	Sequence 7068, Ap
C 19	33.2	11.3	1273	10	US-09-938-842A-3405	Sequence 3405, Ap
C 20	33.2	11.3	2007	10	US-09-887-576-288	Sequence 288, App
C 21	33	11.2	2345	13	US-10-027-632-110060	Sequence 110060,
C 22	33	11.2	2345	13	US-10-027-632-110061	Sequence 110061,
C 23	33	11.2	2345	14	US-10-027-632-110060	Sequence 110060,
C 24	33	11.2	2345	14	US-10-027-632-110061	Sequence 110061,
C 25	33	11.2	14368	13	US-10-017-161-727	Sequence 727, App
C 26	33	11.2	43818	12	US-10-292-798-637	Sequence 637, App
C 27	33	11.2	640681	10	US-09-790-988-1	Sequence 1, Appl1
C 28	32.8	11.2	458	13	US-10-027-632-291958	Sequence 291958,
C 29	32.8	11.2	458	14	US-10-027-632-291958	Sequence 291958,
C 30	32.8	11.2	202001	9	US-09-734-674-3	Sequence 3, Appl1
C 31	32.8	11.2	202001	15	US-10-274-990-3	Sequence 3, Appl1
C 32	32.6	11.1	2832	13	US-10-101-510-42	Sequence 42, Appl1
C 33	32.2	11.0	658	13	US-10-027-632-282173	Sequence 282173,
C 34	32.2	11.0	658	14	US-10-027-632-282173	Sequence 282173,
C 35	32.2	11.0	895	13	US-10-027-632-121917	Sequence 121917,
C 36	32.2	11.0	895	14	US-10-027-632-121917	Sequence 121917,
C 37	32.2	11.0	2000	10	US-09-938-842A-4674	Sequence 4674, Ap
C 38	32.2	11.0	4110	15	US-10-198-846-9909	Sequence 9909, Ap
C 39	32.2	11.0	3673778	13	US-10-312-841-2	Sequence 2, Appl1
C 40	32	10.9	275	10	US-09-878-574-15709	Sequence 15709, A
C 41	32	10.9	424	15	US-10-102-524-227	Sequence 227, App
C 42	32	10.9	1843	10	US-09-880-107-3412	Sequence 3412, App
C 43	32	10.9	3012	11	US-09-919-039-167	Sequence 167, App
C 44	32	10.9	9834	8	US-08-781-986A-37	Sequence 37, Appl1
C 45	31.8	10.8	418	11	US-09-918-995-3050	Sequence 3050, Ap

ALIGNMENTS

RESULT 1
US-10-001-254-5
Sequence 5, Application US/10001254
Publication NO. US20030049702A1
GENERAL INFORMATION:
APPLICANT: Reed, John C.
APPLICANT: Godzik, Adam
APPLICANT: Pawlowski, Krzysztof
APPLICANT: Fiorentino, Loredana
APPLICANT: Lee, Sug Hyung
APPLICANT: Roth, Wilfred
APPLICANT: Steiner-Liewen, Frank
TITLE OF INVENTION: No. US20030049702A1e1 Death Domain Proteins
FILE REFERENCE: P-Id 5037
CURRENT APPLICATION NUMBER: US/10/001,254
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/301,889
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/715,893
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 294
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: CDS
LOCATION: (1 ... (294)
US-10-001-254-5
Query Match 100.0%; Score 294; DB 15; Length 294;
Best Local Similarity 100.0%; Pred. No. 6.4e-75;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACATATGTGGCTGCCTCATATTGAGACTATTAGAGAGCTGTGACATTTTATTTATTCCT 60
Db 1 ACATATGTGGCTGCCTCATATTGAGACTATTAGAGAGCTGTGACATTTTATTTATTCCT 60

Oy	6	CAAGAAGATGAAAGAAAGTTAGCTGAGCTATTAAAAACATCTGGGATGATAGATAC	120
Dd	61	CMAAGAGATGAAGAAAGTTAGCTGACTATTAAAAACCATCTGGGATGATAGATAC	120
Oy	121	AATCAGTTCACATAAAGAGATTGAAAGCATTACTTCAAACGTGAAAAGTCCACTTCT	180
Dd	121	AATCAGTTCACATAAAGAGATTGAAAGCATTACTTCAAACGTGAAAAGTCCACTTCT	180
Oy	181	GAAATTACTGTTTGACTCGGGGCACACAATTTGACACAGTTGGTAGTCTTGATGATCTTTTG	240
Dd	181	GAAATTACTGTTTGACTCGGGGCACACAATTTGACACAGTTGGTAGTCTTGATGATCTTTTG	240
Oy	241	ATCCAAAATAATTTTTTGCTCCCTGGAGATCTTTTGCTCCAGATGCTGTTCCC	294
Dd	241	ATCCAAAATAATTTTTTGCTCCCTGGAGATCTTTTGCTCCAGATGCTGTTCCC	294

```

RESULT 2
US-10-001-254-15
Sequence 15, Application US/10001254
Publication No. US20030049702A1
GENERAL INFORMATION:
APPLICANT: Reed, John C.
APPLICANT: Godzik, Adam
APPLICANT: Pawlowski, Krzysztof
APPLICANT: Fiorentino, Loredana
APPLICANT: Lee, Sug Hyung
APPLICANT: Roth, Wilfred
APPLICANT: Stemmer-Wilfeden, Frank
TITLE OR INVENTION: No. US20030049702A1e1 Death Domain Proteins
FILE REFERENCE: P-LJ 5037
CURRENT APPLICATION NUMBER: US/10/001,254
CURRENT FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/301,889
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/715,893
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 1383
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1380)
US-10-001-254-15

```

Query Match	100.0%	Score 294	DB 15	Length 1383
Beet Local Similarity	100.0%	Pred. No. 1.3e-74		
Matches 294	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	ACATATGTCGCTGCCTCAATGTTGAGACAAATAGGAACCTGTCAGATTATATGATCCT	60	
Db	25	ACATATGTCGCTGCCTCAATGTTGAGACAAATAGGAACCTGTCAGATTATATGATCCT	84	
Qy	61	CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAAACATCTGGTATGATATAC	120	
Db	85	CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAAACATCTGGTATGATATAC	144	
Qy	121	AATCAGTTTCACATTAAGAGATTTGAAGCATTCCTCAACTGGAAAAAGTCCCACTTCT	180	
Db	145	AATCAGTTTCACATTAAGAGATTTGAAGCATTCCTCAACTGGAAAAAGTCCCACTTCT	204	
Qy	181	GAATTACTGTTGACTGGGGCACACAAATTGACAGTTGGTGTGATCTTTTGG	240	
Db	205	GAATTACTGTTGACTGGGGCACACAAATTGACAGTTGGTGTGATCTTTTGG	264	
Qy	241	ATCCAAATGATTTTTTGGCTCTGGGAGCTTTTGGCTCCAGATGCTGTCC	294	
Db	265	ATCCAAATGATTTTTTGGCTCTGGGAGCTTTTGGCTCCAGATGCTGTGTCC	318	

```

RESULT 3
US-09-966-451-3
; Sequence 3, Application US/09966451
; Publication No. US20030087856A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freiler
; TITLE OF INVENTION: ANTISENSE MODULATION OF IL-1 RECEPTOR-ASSOCIATED KINASE-4 EXPRESSION
; FILE REFERENCE: RFS-0324
; CURRENT APPLICATION NUMBER: US/09/966,451
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 88
; SEQ ID NO 3
; LENGTH: 2817
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)... (1432)
; US-09-966-451-3

```

	Query Match	Similarity	Score	DB	Length	Best Local	Mismatches	Indels	Gaps
	Matches	Conservative	0;	294;	2817;	100.0%;	0;	0;	0
			0;	1.9e-74;		100.0%;	0;	0;	0
QY	1	ACATATGTGCGCTGCTCAATGTTGACATAATTAGAAAGCTGTCAAGATTTTATTGATCT	60						
DB	74	ACATATGTGCGCTGCTCTCAATGTTGACATAATTAGAAAGCTGTCAAGATTTTATTGATCT	133						
QY	61	CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAAACATCTGTGATGATGATATAC	120						
DB	134	CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAAACATCTGTGATGATGATATAC	193						
QY	121	AATCAGTTTCAATPAGAGATTTGAGCATTACTTCAAACTGGAAAAAGTCCCACTCT	180						
DB	194	AATCAGTTTCAATPAGAGATTTGAGCATTACTTCAAACTGGAAAAAGTCCCACTCT	253						
QY	181	GAATTACTGTTGACTGGGACACCACAAAATTGACAGTTGTGATCTTGTGATCTTTTG	240						
DB	254	GAATTACTGTTGACTGGGACACCACAAAATTGACAGTTGTGATCTTGTGATCTTTTG	313						
QY	241	AATCCAAATGAATTTTGTCTCTGCCAGTCTTTTGTCTCCAGATGCTGTTCCC	294						
DB	314	AATCCAAATGAATTTTGTCTCTGCCAGTCTTTTGTCTCCAGATGCTGTTCCC	367						

```

/ RESULT 4
/ US-10-001-254-27
/ Sequence 27, Application US/10001254
/ Publication No. US20030049702A1
/ GENERAL INFORMATION:
/ APPLICANT: Reed, John C.
/ APPLICANT: Godzik, Adam
/ APPLICANT: Pawlowski, Krzysztof
/ APPLICANT: Fiorentino, Lorendana
/ APPLICANT: lee, Sung Hyung
/ APPLICANT: Roth, Wilfried
/ APPLICANT: Stehner-Nielsen, Frank
/ TITLE OF INVENTION: No. US20030049702A1e1 Deach Domain Proteins
/ FILE REFERENCE: P-IJ 5037
/ CURRENT APPLICATION NUMBER: US/10/001,254
/ CURRENT FILING DATE: 2001-11-15
/ PRIOR APPLICATION NUMBER: 60/301,889
/ PRIOR FILING DATE: 2001-06-29
/ PRIOR APPLICATION NUMBER: 09/115,893
/ PRIOR FILING DATE: 2000-11-17
/ NUMBER OF SEQ ID NOS: 62
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 27
/ LENGTH: 2817
/ TYPE: DNA
/ ORGANISM: Homo sapien

```


FEATURE:
NAME/KEY: CDS
LOCATION: (50) ... (1429)
US-10-001-254-27

Query Match 100.0%; Score 294; DB 15; Length 2817;
Best Local Similarity 100.0%; Pred. No. 1.9e-74;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACATATGTGGCGCTCCATAGTTGAGCTAATTAGGAAGCTGTCAGATTTTATGATCT 60
DB 74 ACATATGTGGCGCTCCATAGTTGAGCTAATTAGGAAGCTGTCAGATTTTATGATCT 133
QY 61 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTTAAAAACCATCTGGTGAATAGATAC 120
DB 134 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTTAAAAACCATCTGGTGAATAGATAC 193
QY 121 AATCAGTTTGCATAGAGAGATTGGAAGCATTACTTCAAACTGGAAAAAGTCCACTTCT 180
DB 194 AATCAGTTTGCATAGAGAGATTGGAAGCATTACTTCAAACTGGAAAAAGTCCACTTCT 253
QY 181 GAATTAATGTTGCTGGGGGACCAAAATTTGACAGTGTATCTTTGGATCTTTTG 240
DB 254 GAATTAATGTTGCTGGGGGACCAAAATTTGACAGTGTATCTTTGGATCTTTTG 313
QY 241 ATCCAAATGAAATTTTGTCTCTGCGAGTCTTTGCTCCAGATGCTGTTCC 294
DB 314 ATCCAAATGAAATTTTGTCTCTGCGAGTCTTTGCTCCAGATGCTGTTCC 367

RESULT 5
US-09-759-595-2
Sequence 2, Application US/09759595
Publication No. US2003005916A1

GENERAL INFORMATION:
APPLICANT: Wesche, Holger
APPLICANT: Li, Shyun
APPLICANT: Tularik Inc.
TITLE OF INVENTION: IRAK-4: Compositions and Methods of Use
FILE REFERENCE: 018781-003910US
CURRENT APPLICATION NUMBER: US/09/759,595
CURRENT FILING DATE: 2001-01-13
PRIOR APPLICATION NUMBER: US 60/176,395
PRIORITY FILING DATE: 2000-01-13
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1383
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human IL-1 receptor-associated kinase 4 (IRAK-4)
OTHER INFORMATION: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: (1) ... (1383)
OTHER INFORMATION: human IRAK-4
US-09-759-595-2

Query Match 99.5%; Score 292.4; DB 11; Length 1383;
Best Local Similarity 99.7%; Pred. No. 3.9e-74;
Matches 293; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACATATGTGGCGCTCCATAGTTGAGCTAATTAGGAAGCTGTCAGATTTTATGATCT 60
DB 25 ACATATGTGGCGCTCCATAGTTGAGCTAATTAGGAAGCTGTCAGATTTTATGATCT 84
QY 61 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTTAAAAACCATCTGGTGAATAGATAC 120
DB 85 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTTAAAAACCATCTGGTGAATAGATAC 144
QY 121 AATCAGTTTGCATAGAGAGATTGGAAGCATTACTTCAAACTGGAAAAAGTCCACTTCT 180

DB 145 AATCAGTTTGCATAGAGAGATTGGAAGCATTACTTCAAACTGGAAAAAGTCCACTTCT 204
QY 181 GAATTAATGTTGCTGGGGGACCAAAATTTGACAGTGTATCTTTGGATCTTTTG 240
DB 205 GAATTAATGTTGCTGGGGGACCAAAATTTGACAGTGTATCTTTGGATCTTTTG 264
QY 241 ATCCAAATGAAATTTTGTCTCTGCGAGTCTTTGCTCCAGATGCTGTTCC 294
DB 265 ATCCAAATGAAATTTTGTCTCTGCGAGTCTTTGCTCCAGATGCTGTTCC 318

RESULT 6
US-09-833-790-149

Sequence 149, Application US/09833790
Patent No. US2002006828A1
GENERAL INFORMATION:
APPLICANT: Lodes, Michael J.
APPLICANT: Wang, Tonglong
APPLICANT: Secrist, Heather
APPLICANT: Mohamath, Raodoh
APPLICANT: Indirias, Carol Y.
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.512
CURRENT APPLICATION NUMBER: US/09/833,790
CURRENT FILING DATE: 2001-04-11
NUMBER OF SEQ ID NOS: 440
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 149
LENGTH: 501
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1) ... (501)
OTHER INFORMATION: n = A,T,C or G
US-09-833-790-149

Query Match 82.2%; Score 241.6; DB 9; Length 501;
Best Local Similarity 97.3%; Pred. No. 1.1e-59;
Matches 288; Conservative 0; Mismatches 4; Indels 4; Gaps 4;

QY 1 ACATATGTGGCGCTCCATAGTTGAGCTAATTAGGAAGCTGTCAGATTTTATGATCT 60
DB 45 ACATATGTGGCGCTCCATAGTTGAGCTAATTAGGAAGCTGTCAGATTTTATGATCT 104
QY 61 CAAGAAGATGGAAGAAGTTAGCTGTAGCTAATTAAAAACCATCTGGTGAATAGATAC 120
DB 105 CAAGAAGATGGAAGAAGTTAGCTGTAGCTAATTAAAAACCATCTGGTGAATAGATAC 164
QY 121 AATC-AGTTTGCATAGAGAGATTGGAAGCATTACTTCAAACTGGAAAAAGTCCACTTCT 179
DB 165 AATCAGTTTGCATAGAGAGATTGGAAGCATTACTTCAAACTGGAAAAAGTCCACTTCT 223
QY 180 TGAATTAATGTTGCTGGGGGACCAAAATTTGACAGTGTATCTTTGGATCTTTTG 238
DB 224 TGAATTAATGTTGCTGGGGGACCAAAATTTGACAGTGTATCTTTGGATCTTTTG 283
QY 239 TGAATTAATGAAATTTTGTCTCTGCGAGTCTTTGCTCCAGATGCTGTTCC 294
DB 284 TGAATTAATGAAATTTTGTCTCTGCGAGTCTTTGCTCCAGATGCTGTTCC 338

RESULT 7
US-09-759-595-4

Sequence 4, Application US/09759595
Publication No. US2003005916A1
GENERAL INFORMATION:
APPLICANT: Wesche, Holger
APPLICANT: Li, Shyun
APPLICANT: Tularik Inc.
TITLE OF INVENTION: IRAK-4: Compositions and Methods of Use

FILE REFERENCE: 018781-003910US
CURRENT APPLICATION NUMBER: US/09/759,595
CURRENT FILING DATE: 2001-01-13
PRIOR APPLICATION NUMBER: US 60/176,395
PRIOR FILING DATE: 2000-01-13
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 1542
TYPE: DNA
ORGANISM: Mus sp.
FEATURE:
OTHER INFORMATION: murine IL-1 receptor-associated kinase 4 (IRAK-4)
FEATURE:
OTHER INFORMATION: CDNA
NAME/KEY: CDS
LOCATION: (163)..(1542)
OTHER INFORMATION: murine IRAK-4
US-09-759-595-4

Query Match 69.5%; Score 204.4; DB 11; Length 1542;
Best Local Similarity 81.0%; Pred. No. 1.1e-48;
Matches 238; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1 ACATATGTCGCTGCTCAATGTTGACTAATTAGAGCTGTGAGATTTTATGATCCT 60
DB 187 ACATATATGCGCAACTTAATGTTGGGATCTTGAAGAGCTGTGAGATTTTATGATCCT 246
QY 61 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAACATCTGTGTATGATATAC 120
DB 247 CAAGAAGGTTGGAAGAATTTAGCACTATCAAAAAGCCCTCCGCGACGACATAC 306
QY 121 AATCAGTTTACATAAGAGATTGGAAGCTTCAACTGGAAGAAAGTCCCACTCT 180
DB 307 AATCAGTTTCAATTAAGAGATTGGAAGCTTCAACTGGAAGAAAGTCCCACTCT 366
QY 181 GAATTAATGTTGACTGCGGACACACAATTTGACAGTTGTGATCTTTTGT 240
DB 367 GAATCTGTTTACTGCGGACACACAAGTTGAGCTTGTGATCTTACTG 426
QY 241 ATCCAAATGAATTTTGTCTGCTGAGATCTTTTGTCTCCCAATGCTGTCC 294
DB 427 GTCCAGATTGAGCTGTTTGTCCCGCCGCACTCTCTGCTCGGATCCGTTCC 480

RESULT 8

US-09-966-451-10
Sequence 10, Application US/0966451
Publication No. US20030087856A1
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Susan M. Freier
TITLE OF INVENTION: ANTISENSE MODULATION OF IL-1 RECEPTOR-ASSOCIATED KINASE-4 EXPRESSION
CURRENT APPLICATION NUMBER: US/09/966,451
CURRENT FILING DATE: 2001-09-28
NUMBER OF SEQ ID NOS: 88
SEQ ID NO 10
LENGTH: 31000
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-09-966-451-10

Query Match 51.6%; Score 151.8; DB 11; Length 31000;
Best Local Similarity 95.7%; Pred. No. 7.4e-33;
Matches 156; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 123 TCAGTTTCACATAGAGATTGAAGCTTCAAACTGGAAGAAAGTCCCACTCTGA 182
DB 12953 TTACTTACTTTTAAGAGATTGAAGCTTCAAACTGGAAGAAAGTCCCACTCTGA 13012
QY 183 ATTACTGTTTGAAGCTGCGGACCAAAATTCACAGTTGTGATCTTGTGATCTTTTGA 242

DB 13013 ATTACTGTTTGAAGCTGCGGACCAAAATTCACAGTTGTGATCTTGTGATCTTTTGA 13072
QY 243 CCAAAATGAATTTTGTCTCCCGAGAGTCTTTGTCTCCCAAT 285
DB 13073 CCAAAATGAATTTTGTCTCCCGAGAGTCTTTGTCTCCCAAT 13115

RESULT 9

US-10-001-254-25
Sequence 25, Application US/10001254
Publication No. US20030049702A1
GENERAL INFORMATION:
APPLICANT: Reed, John C.
APPLICANT: Godzik, Adam
APPLICANT: Pawlowski, Krzysztof
APPLICANT: Fiorentino, Loredana
APPLICANT: Lee, Sung Hyung
APPLICANT: Roth, Wilfried
APPLICANT: Stenmer-Liawen, Frank
TITLE OF INVENTION: No. US20030049702A1e1 Death Domain Proteins
FILE REFERENCE: P-IL 5037
CURRENT APPLICATION NUMBER: US/10/001,254
CURRENT FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/301,889
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/715,893
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 211
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(177)
US-10-001-254-25

Query Match 46.8%; Score 137.6; DB 15; Length 211;
Best Local Similarity 97.2%; Pred. No. 8.5e-30;
Matches 140; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACATATGTCGCTGCTCAATGTTGACTAATTAGAGCTGTGAGATTTTATGATCCT 60
DB 25 ACATATGTCGCTGCTCAATGTTGACTAATTAGAGCTGTGAGATTTTATGATCCT 84
QY 61 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAACATCTGTGTATGATATAC 120
DB 85 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAACATCTGTGTATGATATAC 144
QY 121 AATCAGTTTACATAAGAGATT 144
DB 145 AATCAGTTTACATAAGATGCTGT 168

RESULT 10

US-09-887-586A-51
Sequence 51, Application US/09887586A
Patent No. US20020094556A1
GENERAL INFORMATION:
APPLICANT: Chappell, Joseph
APPLICANT: No. US20020094556A11, Joseph P.
APPLICANT: Starks, Courtney M.
APPLICANT: Manna, Kathleen R.
TITLE OF INVENTION: SYNTHASES
FILE REFERENCE: 07678-025001
CURRENT APPLICATION NUMBER: US/09/887,586A
CURRENT FILING DATE: 2001-06-22
PRIOR APPLICATION NUMBER: 09/398,395
PRIOR FILING DATE: 1999-09-17
PRIOR APPLICATION NUMBER: 60/130,628
PRIOR FILING DATE: 1999-04-22

```

RESULT 13
US-09-938-842A-4518
; Sequence 4518, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:

```

```
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO: 4518
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4518
```

```
Query Match 12.2%; Score 36; DB 10; Length 2000;
Best Local Similarity 58.3%; Pred. No. 5.8;
Matches 63; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
```

```
QY 111 TGAATGATACATGAGTTTCATAGAGATTGAGAGCATTAATGAGAAAG 170
DB 1374 TTATGGAAGTAATGATTTTATTTAGAGATTTTAAATTAATGATATTTAAATG 1433
QY 171 TCCCATCTTGAAATTAAGTTGACTGGGCGACCAAAATGCAAGT 218
DB 1434 TACAATATATCATTTATTTTGAAGATTTACATTAATTTGCAAGT 1481
```

RESULT 14

```
US-09-960-858-2
; Sequence 2, Application US/09960858
; Publication No. US20030138777A1
; GENERAL INFORMATION:
; APPLICANT: Evans, Glen
; TITLE OF INVENTION: NANOMACHINE COMPOSITIONS AND METHODS OF
; FILE REFERENCE: P-EA 4974
; CURRENT APPLICATION NUMBER: US/09/960,858
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 2
; LENGTH: 22684
; TYPE: DNA
; ORGANISM: M. genitalium
US-09-960-858-2
```

```
Query Match 11.5%; Score 33.8; DB 13; Length 22684;
Best Local Similarity 53.4%; Pred. No. 80;
Matches 71; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
```

```
QY 143 TTGAAGCATTAAGTCAAAAGTCCCACTTGAATTAAGTTGACTGGGCA 202
DB 17554 TTGAAGCAATGTTGAAACCAAAACCAATCTGATGATTAATTTGACAGATTG 17613
QY 203 CCACAATTTGACAGTTGATCTGTGATCTTTGATGCAAAATGATTTTGTCTC 262
DB 17614 TTAATCGTACTCAATTAATCTTAATGTTGATCTGAAACCAAAACCAAACTATTTGAGAAG 17673
QY 263 CTGCGAGTCTTTT 275
DB 17674 CTACAGCTTTAT 17686
```

```
RESULT 15
US-09-960-870-2
```

```
; Sequence 2, Application US/09960870
; Publication No. US20030134281A1
; GENERAL INFORMATION:
; APPLICANT: Evans, Glen
; TITLE OF INVENTION: NANOMACHINE COMPOSITIONS AND METHODS OF
; TITLE OF INVENTION: USE
; FILE REFERENCE: P-EA 4738
; CURRENT APPLICATION NUMBER: US/09/960,870
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 2
; LENGTH: 22684
; TYPE: DNA
; ORGANISM: M. genitalium
US-09-960-870-2
```

```
Query Match 11.5%; Score 33.8; DB 13; Length 22684;
Best Local Similarity 53.4%; Pred. No. 80;
Matches 71; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
```

```
QY 143 TTGAAGCATTAAGTCAAAAGTCCCACTTGAATTAAGTTGACTGGGCA 202
DB 17554 TTGAAGCAATGTTGAAACCAAAACCAATCTGATGATTAATTTGACAGATTG 17613
QY 203 CCACAATTTGACAGTTGATCTGTGATCTTTGATGCAAAATGATTTTGTCTC 262
DB 17614 TTAATCGTACTCAATTAATCTTAATGTTGATCTGAAACCAAAACCAAACTATTTGAGAAG 17673
QY 263 CTGCGAGTCTTTT 275
DB 17674 CTACAGCTTTAT 17686
```

```
Search completed: January 18, 2004, 01:19:33
Job time : 218 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 21:47:26 ; Search time 1514 Seconds

(without alignments)
7944.142 Million cell updates/sec

Title: US-10-001-254-5

Perfect score: 294
Sequence: 1 acatcgtcgcgcgcctcaaa.....tgcctccagatgctgtccc 294

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

```

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_ov:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_higo_hum:*
40: em_higo_mus:*
41: em_higo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	294	100.0	294	6 AX431296	AX431296 Sequence
2	294	100.0	833	6 AR223870	AR223870 Sequence
3	294	100.0	1383	6 AX431306	AX431306 Sequence
4	294	100.0	2817	6 AX431318	AX431318 Sequence
5	294	100.0	2817	6 AF155118	AF155118 Homo sapi
6	294	100.0	2820	6 AK000528	AK000528 Homo sapi
7	292.4	99.5	1383	6 AX196260	AX196260 Sequence
8	292.4	99.5	1383	6 AF445802	AF445802 Homo sapi
9	292.4	99.5	1629	6 BC013316	BC013316 Homo sapi
10	241.6	82.2	501	6 AX321132	AX321132 Sequence
11	204.4	69.5	1542	6 AX196262	AX196262 Sequence
12	204.4	69.5	2431	10 AF445803	AF445803 Mus muscu
13	204.4	69.5	2431	10 BC051676	BC051676 Mus muscu
14	151.8	51.6	3303	9 AY186092	AY186092 Homo sapi
15	151.8	51.6	118572	9 AC093012	AC093012 Homo sapi
16	151.8	51.6	165868	2 AC021719	AC021719 Homo sapi
17	151.8	51.6	309582	2 AC025567	AC025567 Homo sapi
18	137.6	46.8	221	6 AX431316	AX431316 Sequence
19	137.6	46.8	2213	6 BD155790	BD155790 Primer to
20	137.6	46.8	2213	6 AK027301	AK027301 Homo sapi
21	110	37.4	264657	2 AC129390	AC129390 Rattus no
22	94.8	32.2	1719	5 BC045381	BC045381 Dario rer
23	61.4	20.9	244696	2 AC123081	AC123081 Rattus no
24	61.4	20.9	252400	2 AC094859	AC094859 Rattus no
25	54.2	18.4	88858	11 BV075713	BV075713 S212P6044
26	46.8	15.9	110000	2 AC134788_3	Continuation 14 of
27	46.8	15.9	110000	2 AC103663_1	Continuation 12 of
28	43	14.6	9990	1 AE010646	AE010646 Fusobacte
29	42.8	14.6	574	11 BV034947	BV034947 S212P6037
30	42.6	14.5	1694	9 HUMTMP6	AF005361 Homo sapi
31	42.6	14.5	2164	9 BC047409	BC047409 Homo sapi
32	42.4	14.4	2000	8 AF279456	AF279456 Lycopersi
33	41.6	14.1	68858	9 AL139425	AL139425 Human DNA
34	41.6	14.1	185158	9 AC008732	AC008732 Homo sapi
35	41	13.9	160851	5 AL954333	AL954333 zebrafish
36	40.8	13.9	943	11 CNS061ST	AL400707 T7 end of
37	40.8	13.9	980	11 CNS061K2	AL400392 T3 end of
38	40.8	13.9	6095	8 SEA39550	AJ439550 Saccharom
39	40.4	13.7	1879	6 AR183920	AR183920 Sequence
40	40.4	13.7	1879	8 AF035630	AF035630 Lycopersi
41	40.4	13.7	2003	8 AF279454	AF279454 Lycopersi
42	40.4	13.7	2018	8 AF279453	AF279453 Lycopersi
43	40.4	13.7	2024	6 AR183921	AR183921 Sequence
44	40.4	13.7	2024	6 AR240709	AR240709 Sequence
45	40.4	13.7	2024	6 AR266984	AR266984 Sequence

ALIGNMENTS

RESULT 1	AX431296	294 bp	DNA	linear	PAT 28-JUN-2002
LOCUS	AX431296				
DEFINITION	Sequence 5 from Patent WO0240680.				
ACCESSION	AX431296				
VERSION	AX431296.1	GI:21656165			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE					
AUTHORS	Pawlowski, K., Fiorentino, L., Godzik, A., Lee, S.H., Reed, J.C., Roth, W. and Steiner-Jewen, F.				
TITLE	Novel death domain proteins				

FEATURES	source	Location/Qualifiers
JOURNAL	Patient: WO 0240680-A 5 23-MAY-2002;	
FEATURES	BURNHAM INST (US)	
source	1. .294	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	<1. .>294	
	/note="unamed protein product"	
	/codon_start=1	
	/protein_id="CAD37274.1"	
	/db_xref="GI:21656166"	
	/translation="HYVACLVNGLIRKLSDPFDPEGKKILAVIKRPSGDDRNQCH IRREFALQTKSPTESELFDMGTCTVGDVLDLLIQNEFFAPASILLPDVAF"	
CDS		
BASE COUNT	84 a 55 c 62 g 93 t	
ORIGIN		
Query Match	100.0%; Score 294; DB 6; Length 294;	
Best Local Similarity	100.0%; Pred. No. 1.4e-64;	
Matches 294; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Db	1 ACATATGTCGCGTGCCTCAATGTTCGACTAATTAGAACTGTGAGATTTTATTCATCCT 60	
Qy	1 ACATATGTCGCGTGCCTCAATGTTCGACTAATTAGAACTGTGAGATTTTATTCATCCT 60	
Db	1 ACATATGTCGCGTGCCTCAATGTTCGACTAATTAGAACTGTGAGATTTTATTCATCCT 60	
Qy	61 CAGAAGAGATGGAAGAGTTCGTGATCTATTTAAAAACCATCTGGTGTGATGATGATAC 120	
Db	61 CAGAAGAGATGGAAGAGTTCGTGATCTATTTAAAAACCATCTGGTGTGATGATGATAC 120	
Qy	121 AATCAGTTTCATATGAGAGATTTGAAACATTTACTTCAACCTGGAAAAAGTCCACTCT 180	
Db	121 AATCAGTTTCATATGAGAGATTTGAAACATTTACTTCAACCTGGAAAAAGTCCACTCT 180	
Qy	181 GAATTACTGTTTGACTGCGGACCAACAATTGACAGTTGGTGTGATCTTGTGATCTTTTG 240	
Db	181 GAATTACTGTTTGACTGCGGACCAACAATTGACAGTTGGTGTGATCTTGTGATCTTTTG 240	
Qy	241 ATCCAAATGAATTTTTCCTGCTCGAGTCTTTTGTCTCCAGATGCTGTTCCC 294	
Db	241 ATCCAAATGAATTTTTCCTGCTCGAGTCTTTTGTCTCCAGATGCTGTTCCC 294	
RESULT 2	AR223870 833 bp DNA linear PAT 26-SEP-2002	
LOCUS	AR223870	
DEFINITION	Sequence 10 from patent US 6440663.	
ACCESSION	AR223870	
VERSION	AR223870.1 GI:23332452	
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	Unclassified.	
AUTHORS	1 (bases 1 to 833)	
	Scanlan,M.J., Stockert,E., Chen,Y.-T., Old,L.J., Jager,E. and	
	Knuth,A.	
	Renal cancer associated antigens and uses therefor	
	Patent: US 6440663-A 10 27-AUG-2002;	
TITLE	Location/Qualifiers	
JOURNAL	1. 833	
FEATURES	/organism="unknown"	
BASE COUNT	273 a 155 c 179 g 226 t	
ORIGIN		
Query Match	100.0%; Score 294; DB 6; Length 833;	
Best Local Similarity	100.0%; Pred. No. 1.3e-64;	
Matches 294; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Db	1 ACATATGTCGCGTGCCTCAATGTTCGACTAATTAGAACTGTGAGATTTTATTCATCCT 60	
Qy	1 ACATATGTCGCGTGCCTCAATGTTCGACTAATTAGAACTGTGAGATTTTATTCATCCT 60	
Db	74 ACATATGTCGCGTGCCTCAATGTTCGACTAATTAGAACTGTGAGATTTTATTCATCCT 133	
Qy	61 CAGAAGAGATGGAAGAGTTCGTGATCTATTTAAAAACCATCTGGTGTGATGATGATAC 120	

Db	134	CAGAAGAGTGAAGAGTAGTACGTGTAGTATTAAAAACATCTGGTGTAGTAGATAC	193
Qy	121	AATCAGTTTCACATAGAGATTGGAAGCATTTCTAAACTGGAAAAATGCCACTTCT	180
Db	194	AATCAGTTTCACATAGAGATTGGAAGCATTTCTAAACTGGAAAAATGCCACTTCT	253
Qy	181	GAATTACGTTTGAAGCTGGGGGACCAACAAATTGACAGTTGGTATCTTGTGATCTTTTG	240
Db	254	GAATTACGTTTGAAGCTGGGGGACCAACAAATTGACAGTTGGTATCTTGTGATCTTTTG	313
Qy	241	ATCCAAATGAAATTTTGTCTCTCGAGTCCTTTTGTCTCCAGATGCTGTTCC	294
Db	314	ATCCAAATGAAATTTTGTCTCTCGAGTCCTTTTGTCTCCAGATGCTGTTCC	367
RESULT 3			
AX431306	1383 bp	DNA	linear
LOCUS	Sequence 15 from Patent WO0240680.		
DEFINITION	AX431306		
ACCESSION	AX431306.1	GI:21656175	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euleleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 Pawlowski, K., Fiorentino, L., Godzik, A., Lee, S. H., Reed, J. C., Roch, W., and Stenmer-Liawen, F.		
JOURNAL	Novel death domain proteins		
FEATURES	Patent: WO 0240680-A 15 23-May-2002;		
source	BURNHAM INST (US)		
	Location/Qualifiers		
	1..1383		
CDS	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	1..1383		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD37279.1"		
	/db_xref="GI:21656176"		
	/translation="MNKPIPTSPYVRCIANGILIRKSDPIIDPQGMKKLVAIKKPSG		
	DDRNOGHIREFEALDTGKSPYSSELFPDGGTTCYGDVLDLIONEFPAPASLLP		
	DAVRKNTLSEKEATITVOQKQMFCKDRLMTLPVNLQSYMPDSSPENSLELV		
	SDTFHFSPFELKNTVNNNDERIISVGKMGSGGVYKGVNNTVAVKLLAAM		
	VDITTEELKOQFDEIKVMACQHEMLVELLGSDDGDLVYVYVNGVSLDRISC		
	LDGPIPLSMHRCKIAGGAANGINFEHNHHRDISANILDEAFPAKISDGLAR		
	ASERPAQTVMSRIVGTTFAYMAPALGELITPKSDIYSFGVILLITTG.PAVDEHXY		
	PQLLIDKELEIDEKTIETDIIDKKNADOSTSEVANYSAVSCLEHKRKRPRDIKXY		
	PQLLIDKELEIDEKTIETDIIDKKNADOSTSEVANYSAVSCLEHKRKRPRDIKXY		
	QQLQLEMTAS"		
BASE COUNT	463 a	243 c	283 g
ORIGIN			394 t
Query Match	100.0%;	Score 294;	DB 6;
Best Local Similarity	100.0%;	Pred. No. 1.2e-64;	Length 1383;
Matches	294;	Conservative 0;	Mismatches 0;
		Indels 0;	Gaps 0;
Qy	1	ACATATGCGCGTGCCTCAATGTTGGACATATTGAGAAAGCTGCAGATTTTATGATCCT	60
Db	25	ACATATGCGCGTGCCTCAATGTTGGACATATTGAGAAAGCTGCAGATTTTATGATCCT	84
Qy	61	CAAGAAGANTGGAAGAAGTGTAGCTGTAGCTATTAAAAAACCATCTGTGTGATGATGATAC	120
Db	85	CAAGAAGANTGGAAGAAGTGTAGCTGTAGCTATTAAAAAACCATCTGTGTGATGATGATAC	144
Qy	121	AATCAGTTTCACATAGAGATTGGAAGCATTTCTCAAACGTGAAAAAGTCCACTTCT	180
Db	145	AATCAGTTTCACATAGAGATTGGAAGCATTTCTCAAACGTGAAAAAGTCCACTTCT	204
Qy	181	GAATTACGTTTGAAGCTGGGGGACCAACAAATTGACAGTTGGTATCTTGTGATCTTTTG	240
Db	205	GAATTACGTTTGAAGCTGGGGGACCAACAAATTGACAGTTGGTATCTTGTGATCTTTTG	264

Qy 241 ATCCAAATGATTTTCTCTCTCGAGTCTTTTGTCCGAGATGCTGTTCC 294
Db 265 ATCCAAATGATTTTCTCTCTCGAGTCTTTTGTCCGAGATGCTGTTCC 318

RESULT 4
AX431318 2817 bp DNA linear PAT 28-JUN-2002
LOCUS Sequence 27 from Patent WO0240680.
DEFINITION AX431318
ACCESSION AX431318
VERSION AX431318.1 GI:21656187
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Pawlowski, K., Fiorentino, L., Godzik, A., Lee, S.H., Reed, J.C.,
Roch, W., and Stenmer-Liwen, F.
Novel death domain proteins
Patent: WO 0240680-A 27 23-MAY-2002;
JOURNAL BURNHAM INST (US)

FEATURES
source Location/Qualifiers
1..2817
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
50..1432
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD37285.1"
/db_xref="GI:21656188"
/translation="MNKPIRSTYVRCINVLIRKLSFDIPDQGMKLAIVAIKKPSG
DDRNRGRIHREFALLDTGKSPSELSFDMWTNCTGDIYDLIQLNFFAPSLIP
DAVPRNTPLSKKAITVQKMPFCDDRLTLPVONLQSTMPDSSPENSLELV
SDIRFHSFSELYELKVNNTNDFERPIVSGNMGEGGVYKGVNNTVAVKILAM
VDITBELKQDFDQIKVMAKCOHENIYVELIGSPDGLCLVYVMPNGLDLRLSC
LDGTPPLSMHRCIKIAGANGINFLHNHHRIDIKSANILDEAFPAKISDGLAR
ASEKPAOTWMTSRIVGTTAYMAPALRGELTPKSDIYSFGVLIETITGPAVDEHRE
PQLLDIKERIEDEKTIEDYIDKKNADSTVEAMYSAGSQRHKKNSPDIKYV
HQLDQENTAS"

CDS
BASE COUNT 912 a 547 c 586 g 772 t
ORIGIN

Query Match 100.0%; Score 294; DB 6; Length 2817;
Best Local Similarity 100.0%; Pred. No. 1.1e-64;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACATATGCGCGCTCCATGTTGAGTAAATTAAGAAAGCTGAGATTTATGATCT 60
Db 74 ACATATGCGCGCTCCATGTTGAGTAAATTAAGAAAGCTGAGATTTATGATCT 133

Qy 61 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAACCATCTGTGATGATAGATAC 120
Db 134 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAACCATCTGTGATGATAGATAC 193

Qy 121 AATAGATTTCATTAAGAGATTGAGACATTACTCAACTGGAAGAACGCCACTTCT 180
Db 194 AATAGATTTCATTAAGAGATTGAGACATTACTCAACTGGAAGAACGCCACTTCT 253

Qy 181 GAATTAAGTTGAGTGGGGGACCAACAATTTGACAGAGTTGATCTTTGATCTTTTG 240
Db 254 GAATTAAGTTGAGTGGGGGACCAACAATTTGACAGAGTTGATCTTTGATCTTTTG 313

Qy 241 ATCCAAATGATTTTCTCTCTCGAGTCTTTTGTCCGAGATGCTGTTCC 294
Db 314 ATCCAAATGATTTTCTCTCTCGAGTCTTTTGTCCGAGATGCTGTTCC 367

RESULT 5
AF155118 2817 bp mRNA linear PRI 05-JAN-2000
LOCUS

DEFINITION Homo sapiens putative protein kinase NY-REN-64 antigen mRNA,
complete cds.
ACCESSION AF155118
VERSION AF155118.1 GI:5360130
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 2817)
AUTHORS Scanlan, M.J., Gordon, J.D., Williamson, B., Stockert, E., Bander, N.H.,
Jongeneel, V., Gure, A.O., Jager, D., Jager, E., Knuth, A., Chen, Y.-T.,
and Old, L.J.
Antigens recognized by autologous antibody in patients with
renal-cell carcinoma
Int. J. Cancer 83 (4), 456-464 (1999)

TITLE JOURNAL MEDLINE
99438124
10508479
2 (bases 1 to 2817)
PUBMED

REFERENCE
1 Scanlan, M.J., Gordon, J.D., Williamson, B., Stockert, E., Bander, N.H.,
Jongeneel, V., Gure, A.O., Jager, D., Jager, E., Knuth, A., Chen, Y.-T.,
and Old, L.J.
Direct Submission
Submitted (28-MAY-1999) Ludwig Institute, Sloan-Kettering
Institute, 1275 York Ave, New York, NY 10021, USA

JOURNAL
Institute

FEATURES
source Location/Qualifiers
1..2817
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="1973/10-4"
/cell_type="renal cell carcinoma"
50..1432
/codon_start=1
/product="putative protein kinase NY-REN-64 antigen"
/protein_id="A042884.1"
/db_xref="GI:5360131"
/translation="MNKPIRSTYVRCINVLIRKLSFDIPDQGMKLAIVAIKKPSG
DDRNRGRIHREFALLDTGKSPSELSFDMWTNCTGDIYDLIQLNFFAPSLIP
DAVPRNTPLSKKAITVQKMPFCDDRLTLPVONLQSTMPDSSPENSLELV
SDIRFHSFSELYELKVNNTNDFERPIVSGNMGEGGVYKGVNNTVAVKILAM
VDITBELKQDFDQIKVMAKCOHENIYVELIGSPDGLCLVYVMPNGLDLRLSC
LDGTPPLSMHRCIKIAGANGINFLHNHHRIDIKSANILDEAFPAKISDGLAR
ASEKPAOTWMTSRIVGTTAYMAPALRGELTPKSDIYSFGVLIETITGPAVDEHRE
PQLLDIKERIEDEKTIEDYIDKKNADSTVEAMYSAGSQRHKKNSPDIKYV
HQLDQENTAS"

CDS
BASE COUNT 912 a 547 c 586 g 772 t
ORIGIN

Query Match 100.0%; Score 294; DB 9; Length 2817;
Best Local Similarity 100.0%; Pred. No. 1.1e-64;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACATATGCGCGCTCCATGTTGAGTAAATTAAGAAAGCTGAGATTTATGATCT 60
Db 74 ACATATGCGCGCTCCATGTTGAGTAAATTAAGAAAGCTGAGATTTATGATCT 133

Qy 61 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAACCATCTGTGATGATAGATAC 120
Db 134 CAAGAAGATGGAAGAAGTTAGCTGTAGCTATTAAAAACCATCTGTGATGATAGATAC 193

Qy 121 AATAGATTTCATTAAGAGATTGAGACATTACTCAACTGGAAGAACGCCACTTCT 180
Db 194 AATAGATTTCATTAAGAGATTGAGACATTACTCAACTGGAAGAACGCCACTTCT 253

Qy 181 GAATTAAGTTGAGTGGGGGACCAACAATTTGACAGAGTTGATCTTTGATCTTTTG 240
Db 254 GAATTAAGTTGAGTGGGGGACCAACAATTTGACAGAGTTGATCTTTGATCTTTTG 313

Qy 241 ATCCAAATGATTTTCTCTCTCGAGTCTTTTGTCCGAGATGCTGTTCC 294
Db 314 ATCCAAATGATTTTCTCTCTCGAGTCTTTTGTCCGAGATGCTGTTCC 367

DNA of sea ID#5

RESULT 6
AK000528
LOCUS Homo sapiens cDNA FL20521 f18, clone KAT10395.
DEFINITION AK000528.1 GI:7020683
ACCESSION AK000528
VERSION oligo capping; f18 (full insert sequence).
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Matarabek, K., Kumagai, A., Itakura, S., Yamazaki, M., Tashiro, H.,
Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Nakamura, Y., Isegai, T. and Sugano, S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2820)
Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isegai, T.,
Shibahara, T., Tanaka, T. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
Minato-Ku, Tokyo 108-8639, Japan (E-mail:cdna1@ims.u-tokyo.ac.jp,
Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan. cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction: 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source location/Qualifiers
1..2820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KAT10395"
/cell_line="KATO III"
/cell_type="signet-ring cell carcinoma"
/clone_lib="KAT"
/note="cloning vector pME18SFL3"
30..1412
/note="unnamed protein product"
/codon_start=1
/protein_id="BA91232.1"
/db_xref="GI:7020684"
/translation="MNKPIPTSTYVRCNLVGLIRKLSDFIDPQEGMKLVAIKKPSG
DREYNOFHIRFPAALLOTGKSTSELPDMGTTCVGDVLDLIQNEFPASLLP
DAVPTANTLPSKEALITVOOKMPDCDRTMTVPONLEOSYMPDSSPEKSLSEV
SDTRHFSFSEYELKAVTNNDERPIISVGNKMGEGFQVYVYVNTVYAVKLAAM
VDITTEBLKQDFDEIKVMKCOHENLVELLGFSSDGDCLVYVMPGSLDLRLSC
LDITPPLSWMRCKIAOGANGINFLHENDHHRDKSANILLDEAFKISDFGLAR
ASERPAQVYVTSRIYGTAYMAPEALRGELTPSDISYFQVLLLEITIGLPAVDEHRE
POLLIDIKERIEDEKTIEDYIDKKNMDADSTVEAMYSVASQCLHEKKNKRPDIKYV
QQLQEMTAS"
BASE COUNT 940 a 534 c 579 g 767 t
ORIGIN
Query Match 100.0%; Score 294; DB 9; Length 2820;
Best Local Similarity 100.0%; Pred. No. 1,1e-64;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 174 AATCAGTTTCACATTAAGAGATTGAGATTACTTCAAACTGGAAGAAAGTCCACTTCT 233
QY 181 GAATTACGTTTGAATCGGGGACCAAAATTCACAGTTGATGATCTTGATCTTTG 240
Db 234 GAATTACGTTTGAATCGGGGACCAAAATTCACAGTTGATGATCTTTGATCTTTG 293
QY 241 ATCCAAATGAATTTTGTCTCTCGAGTCTTTGCTCCAGATGCTGTTCC 294
Db 294 ATCCAAATGAATTTTGTCTCTCGAGTCTTTGCTCCAGATGCTGTTCC 347
RESULT 7
AX196260
LOCUS Sequence 2 from Patent WO0151641.
DEFINITION AX196260
ACCESSION AX196260.1 GI:15386462
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Wesche, H. and Li, S.
TITLE 1
JOURNAL Patent: WO 0151641-A 2 19-JUL-2001;
Tularik Inc. (US)
FEATURES
source location/Qualifiers
1..1383
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="human IL-1 receptor-associated kinase 4 (IRAK-4)
cDNA"
1..1383
/note="human IRAK-4"
/codon_start=1
/protein_id="CAC6090.1"
/db_xref="GI:15386463"
/translation="MNKPIPTSTYVRCNLVGLIRKLSDFIDPQEGMKLVAIKKPSG
DREYNOFHIRFPAALLOTGKSTSELPDMGTTCVGDVLDLIQNEFPASLLP
DAVPTANTLPSKEALITVOOKMPDCDRTMTVPONLEOSYMPDSSPEKSLSEV
SDTRHFSFSEYELKAVTNNDERPIISVGNKMGEGFQVYVYVNTVYAVKLAAM
VDITTEBLKQDFDEIKVMKCOHENLVELLGFSSDGDCLVYVMPGSLDLRLSC
LDITPPLSWMRCKIAOGANGINFLHENDHHRDKSANILLDEAFKISDFGLAR
ASERPAQVYVTSRIYGTAYMAPEALRGELTPSDISYFQVLLLEITIGLPAVDEHRE
POLLIDIKERIEDEKTIEDYIDKKNMDADSTVEAMYSVASQCLHEKKNKRPDIKYV
QQLQEMTAS"
BASE COUNT 463 a 244 c 283 g 393 t
ORIGIN
Query Match 99.5%; Score 292.4; DB 6; Length 1383;
Best Local Similarity 99.7%; Pred. No. 3,1e-64;
Matches 293; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 265 ATCCAAATGAATTTTGTCTCGGAGTCTTTGTCTCCAGATGCTGTCC 318

RESULT 8
LOCUS AF445802 1383 bp mRNA linear PRI 20-APR-2002
DEFINITION Homo sapiens interleukin-1 receptor associated kinase 4 (IRAK4)
ACCESSION AF445802
VERSION AF445802.1 GI:20219009
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Li, S., Strelow, A., Fontana, E.J. and Wesche, H.
TITLE IRAK-4: a novel member of the IRAK family with the properties of an IRAK-kinase
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (8), 5567-5572 (2002)
MEDLINE 21957277
PUBMED 11960013
REFERENCE
AUTHORS Suzuki, N., Suzuki, S., Duncan, G.S., Millar, D.G., Wada, T., Matsuo, C., Takada, H., Wakeham, A., Itie, A., Li, S., Penninger, J.M., Wesche, H., Ohashi, P.S., Mak, T.W. and Yeh, W.C.
TITLE Severe impairment of interleukin-1 and Toll-like receptor signalling in mice lacking IRAK-4
JOURNAL Nature 416 (6882), 750-756 (2002)
MEDLINE 21953935
PUBMED 11923871
REFERENCE
AUTHORS Li, S., Strelow, A., Fontana, E.J. and Wesche, H.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-2001) Biology I, Tularex Inc., 2 Corporate Drive, South San Francisco, CA 94080, USA
FEATURES
source
1. .1383
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .1383
/gene="IRAK4"
1. .1383
/gene="IRAK4"
/note="protein kinase"
/codon_start=1
/product="interleukin-1 receptor associated kinase 4"
/protein_id="AA15772.1"
/db_xref="GI:20219010"
/translation="MNKPTPTSTYVRCNLVGLIRKLSDFIDPOEGMKLAIAIKKPSG
DVRNQHFRIRFEALQTKSPSELFPMTGTTCTAGDVLIDLIQNEFPAPASLLP
DAVPTKANTLPSKEAITVOOKMPCDKRLMTPTVONLEOSYMPDSSPENKSLFV
SDRFHSFSEYELKNTNNDERPI SVGNKMGSGFGVYKGYNTVAIVKLAAM
VDITBELKQDFDEIKVAKCOHENLVELLFGSSGDDCLVYVYMPNGSLDRISC
LDGTPPLSMNRCKIAQGAANGINLHEHNRHIDIKSNIILDEAFATPSRGLAR
ASEKFAQTWMTSRIVGTATYMAPALRGRTTPKSDIYSRVVLEIITGCPAYDREH
POLLDIKKEIDEKTEIDYIDKKNNDADSTSVAMYSVAQCLEHKKRPPDIKV
QQLQENTAS"

BASE COUNT 463 a 244 c 283 g 393 t
ORIGIN

Query Match 99.5%; Score 292.4; DB 9; Length 1383;
Best Local Similarity 99.7%; Pred. No. 3.1e-64;
Matches 293; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACATATGTGGCTCTCAATGTTGACTAATTAAGAAGCTGCAGATTTTATGATCCT 60
DB 25 ACATATGTGGCTCTCAATGTTGACTAATTAAGAAGCTGCAGATTTTATGATCCT 84
QY 61 CAAGAAGATGAAGAAGTAGCTGTAGCTATTAATAAACCATCTGCTGATGATAGTAC 120

Db 85 CAAGAAGATGAAGAAGTAGCTGTAGCTATTAATAAACCATCTGCTGATGATAGTAC 144

QY 121 ATCAGTTTCACTAAGAGATTTTGAAGCATTACTTCAACTGGAAGAACGCCACTCT 180
DB 145 ATCAGTTTCACTAAGAGATTTTGAAGCATTACTTCAACTGGAAGAACGCCACTCT 204

QY 181 GAATTAAGTCTTTGATCTGGGACACACAAATTAAGCATTTGATCTTTGATCTTTTG 240
DB 205 GAATTAAGTCTTTGATCTGGGACACACAAATTAAGCATTTGATCTTTGATCTTTTG 264

QY 241 ATCCAAATGAATTTTGTCTCTCGGAGTCTTTGTCTCCAGATGCTGTCC 294
DB 265 ATCCAAATGAATTTTGTCTCTCGGAGTCTTTGTCTCCAGATGCTGTCC 318

RESULT 9
LOCUS BC013316 1629 bp mRNA linear PRI 04-SEP-2001
DEFINITION Homo sapiens, clone MGC:13330 IMAGE:4287014, mRNA, complete cds.
ACCESSION BC013316
VERSION BC013316.1 GI:15426431
KEYWORDS
SOURCE MGC.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA
REMARK
COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapsb-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
Contact: amadansystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Kettelman and Anuradha Madan
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN at: <http://image.llnl.gov>
Series: IRAL Plate: 19 Row: n Column: 24
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7705840.
FEATURES
source
1. .1629
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:13330 IMAGE:4287014"
/tissue_type="Brain, Primitive neuroectodermal"
/clone_lib="NIH_MGC_56"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
71. .1453
/codon_start=1
/product="Unknown (protein for MGC:13330)"
/protein_id="AA113316.1"
/db_xref="GI:15426432"
/translation="MNKPTPTSTYVRCNLVGLIRKLSDFIDPOEGMKLAIAIKKPSG
DVRNQHFRIRFEALQTKSPSELFPMTGTTCTAGDVLIDLIQNEFPAPASLLP
DAVPTKANTLPSKEAITVOOKMPCDKRLMTPTVONLEOSYMPDSSPENKSLFV
SDRFHSFSEYELKNTNNDERPI SVGNKMGSGFGVYKGYNTVAIVKLAAM
VDITBELKQDFDEIKVAKCOHENLVELLFGSSGDDCLVYVYMPNGSLDRISC
LDGTPPLSMNRCKIAQGAANGINLHEHNRHIDIKSNIILDEAFATPSRGLAR
ASEKFAQTWMTSRIVGTATYMAPALRGRTTPKSDIYSRVVLEIITGCPAYDREH
POLLDIKKEIDEKTEIDYIDKKNNDADSTSVAMYSVAQCLEHKKRPPDIKV

BASE COUNT 542 a 291 c 328 g 468 t
ORIGIN
Query Match 99.5%; Score 292.4; DB 9; Length 1629;
Best Local Similarity 99.7%; Pred. No. 3.1e-64;
Matches 293; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OOLQEMTAS"
QY 1 ACATATGCGCGCTGCTCAATGTTGACATTAATAGGAGAGTGCAGATTTTATGATCCT 60
DB 95 ACATATGCGCGCTGCTCAATGTTGACATTAATAGGAGAGTGCAGATTTTATGATCCT 154
QY 61 CAAGAAGATGGAAGAAGTTAGCTGATTAATAAACAATCTGCTGATGATAGATAC 120
DB 155 CAAGAAGATGGAAGAAGTTAGCTGATTAATAAACAATCTGCTGATGATAGATAC 214
QY 121 AATCAGTTTCACATTAAGAGATTTAGAGATTAATCAATGGAAGAAAGTCCACTTCT 180
DB 215 AATCAGTTTCACATTAAGAGATTTAGAGATTAATCAATGGAAGAAAGTCCACTTCT 274
QY 181 GAATTAAGTTTGAATGAGGACACCAAAATGACAGTTGATCTTGTGATCTTTTG 240
DB 275 GAATTAAGTTTGAATGAGGACACCAAAATGACAGTTGATCTTGTGATCTTTTG 334
QY 241 ATCCAAATGAATTTTTTGTCTCTGCGAGTCTTTTGTCTCCAGATGCTGTTCC 294
DB 335 ATCCAAATGAATTTTTTGTCTCTGCGAGTCTTTTGTCTCCAGATGCTGTTCC 388

RESULT 10
AX321132 501 bp DNA linear PAT 15-DEC-2001
LOCUS Sequence 149 from Patent WO0177168.
DEFINITION AX321132
ACCESSION AX321132
VERSION AX321132.1 GI:17904576
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Lodes, M.J., Wang, T., Mohamath, R. and Indirias, C.Y.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0177168-A 149 18-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source location/Qualifiers
1..501
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 166 a 102 c 97 g 134 t 2 others
ORIGIN

Query Match 82.2%; Score 241.6; DB 6; Length 501;
Best Local Similarity 97.3%; Pred. No. 3.2e-51;
Matches 288; Conservative 0; Mismatches 4; Indels 4; Gaps 4;
QY 1 ACATATGCGCGCTGCTCAATGTTGACATTAATAGGAGAGTGCAGATTTTATGATCCT 60
DB 45 ACATATGCGCGCTGCTCAATGTTGACATTAATAGGAGAGTGCAGATTTTATGATCCT 104
QY 61 CAAGAAGATGGAAGAAGTTAGCTGATTAATAAACAATCTGCTGATGATAGATAC 120
DB 105 CAAGAAGATGGAAGAAGTTAGCTGATTAATAAACAATCTGCTGATGATAGATAC 164
QY 121 AATCAGTTTCACATTAAGAGATTTAGAGATTAATCAATGGAAGAAAGTCCACTTCT 179
DB 165 AATCAGTTTCACATTAAGAGATTTAGAGATTAATCAATGGAAGAAAGTCCACTTCT 223
QY 180 TGAATTAAGTTTGAATGAGGACACCAAAATGACAGTTGATCTTGTGATCTTTT 238
DB 224 TGAATTAAGTTTGAATGAGGACACCAAAATGACAGTTGATCTTGTGATCTTTT 283

QY 239 TGATCCAAATGAATTTTTTGTCTCTGCGAGTCTTTTGTCTCCAGATGCTGTTCC 294
DB 284 TGATCCAAATGAATTTTTTGTCTCTGCGAGTCTTTTGTCTCCAGATGCTGTTCC 338

RESULT 11
AX196262 1542 bp DNA linear PAT 28-AUG-2001
LOCUS Sequence 4 from Patent WO0151641.
DEFINITION AX196262
ACCESSION AX196262
VERSION AX196262.1 GI:15386464
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Wesche, H. and Li, S.
TITLE Irak-4: compositions and methods of use
JOURNAL Patent: WO 0151641-A 4 19-JUL-2001;
Tularik Inc. (US)

FEATURES
source location/Qualifiers
1..1542
/organism="Mus sp."
/mol_type="genomic DNA"
/db_xref="taxon:10095"
/note="murine IL-1 receptor-associated kinase 4 (IRAK-4)
cDNA"

CDS

163..1542
/note="murine IRAK-4"
/codon_start=1
/protein_id="CAC6091.1"
/db_xref="GI:15386465"
/translation="MKPLPSTYRINLVGILRLSPIDPQEGMKLAVAIKPSG
DRIHQFIRRFBALLOTGKSPTELDKMTNTTGVLDVLDVLEIPAPATLLP
DAVPQTVSLPREALATVAQTHGPCQEDRTSVMWPLHSCPEPSSPDRSVES
SDTRFHSFHELSKITNNPDOPASAGNMGEQFVVVKGCVNNTIVAKLGM
VEISTEELEKQDFDOI KVMATCOHNLVELLGFSSDSNDLCLVAVMNGSLDLRLSC
LDGPRPLSMHRCVKAOGTANGIRPLRHHNHHRDIKSNILDLDFAKISDFGLAR
ASARLAQVWMSRIYGTATVAPALREIIPKSDIYSGVLLLELITGLAAYDENRE
POLLDIDKEIBDEKTEIDYDERMSADPASYEAMYSAAQSCLHEKKNRPDIAYV
OOLQEMSA"
BASE COUNT 421 a 392 c 423 g 306 t
ORIGIN

Query Match 69.5%; Score 204.4; DB 6; Length 1542;
Best Local Similarity 81.0%; Pred. No. 8.6e-42;
Matches 238; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1 ACATATGCGCGCTGCTCAATGTTGACATTAATAGGAGAGTGCAGATTTTATGATCCT 60
DB 187 ACATATGCGCGCTGCTCAATGTTGACATTAATAGGAGAGTGCAGATTTTATGATCCT 246
QY 61 CAAGAAGATGGAAGAAGTTAGCTGATTAATAAACAATCTGCTGATGATAGATAC 120
DB 247 CAAGAAGATGGAAGAAGTTAGCTGATTAATAAACAATCTGCTGATGATAGATAC 306
QY 121 AATCAGTTTCACATTAAGAGATTTAGAGATTAATCAATGGAAGAAAGTCCACTTCT 180
DB 307 AATCAGTTTCACATTAAGAGATTTAGAGATTAATCAATGGAAGAAAGTCCACTTCT 366
QY 181 GAATTAAGTTTGAATGAGGACACCAAAATGACAGTTGATCTTGTGATCTTTTG 240
DB 367 GAATTAAGTTTGAATGAGGACACCAAAATGACAGTTGATCTTGTGATCTTTTG 426
QY 241 ATCCAAATGAATTTTTTGTCTCTGCGAGTCTTTTGTCTCCAGATGCTGTTCC 294
DB 427 ATCCAAATGAATTTTTTGTCTCTGCGAGTCTTTTGTCTCCAGATGCTGTTCC 480

RESULT 12
AF445803

LOCUS AF445803 1542 bp mRNA linear ROD 20-APR-2002
 DEFINITION Mus musculus interleukin-1 receptor associated kinase 4 (Irak4)
 ACCESSION AF445803
 VERSION AF445803.1 GI:20219011
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1542)
 AUTHORS Li, S., Strelow, A., Fontana, E.J. and Wesche, H.
 TITLE IRAK-4: a novel member of the IRAK family with the properties of an
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (8), 5567-5572 (2002)
 MEDLINE 21957277
 PUBMED 11960013
 REFERENCE 2 (bases 1 to 1542)
 AUTHORS Li, S., Strelow, A., Fontana, E.J. and Wesche, H.
 TITLE Direct Submission
 JOURNAL Submitted (08-NOV-2001) Biology I, Tularek Inc., 2 Corporate Drive,
 South San Francisco, CA 94080, USA
 FEATURES
 Source
 1. 1542
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 gene
 1. 1542
 /gene="Irak4"
 163. 1542
 /gene="Irak4"
 /note="protein kinase"
 /codon_start=1
 /product="interleukin-1 receptor associated kinase 4"
 /protein_id="AA15773.1"
 /db_xref="GI:20219012"
 /translation="MNKPLPSTYIRINLVGILKLSDFIDPOGKKLAIVIKKPSG
 DRRNOPIRREFFALDTGKSPCELLFDGNTTCVGDIVDILVOLEFAPATLLP
 DAVPOTVKSLLPPEAATVNOHGRGCKRDSVWPMKLSHSCPPSSSPDRSVS
 SDTRFHSFTHLEKSTNNEDBOFASAGNMGSGGVYKGVNNTIYAVKLGAM
 VEI STEELKQDFDEIVMAICQEHNVLELFGSSDSNCLVYVNPNSLDRGLSC
 LDGTPPLSMTRCKVQGTANGIRFLHNNHIDIKSANILDKVYNTAKISDFGLAR
 ASARLAQVTMSRIVGTGTAVMAPALRGELTPKSDIYSFGVILLETITGLAAVDENR
 POLIDIKERIEDEKTIEDYTDKMSDADPASYEAMYSASQCLHKKRRPDIAVY
 QQLQESAA"
 BASE COUNT 421 a 392 c 423 g 306 t
 ORIGIN
 Query Match 69.5%; Score 204.4; DB 10; Length 1542;
 Best local similarity 81.0%; Pred. No. 8.6e-42;
 Matches 238; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

RESULT 13
 BC051676
 LOCUS
 DEFINITION Mus musculus interleukin-1 receptor-associated kinase 4, transcript
 variant 1, mRNA (CDNA clone MGC:60994 IMAGE:30017484), complete
 cds.
 ACCESSION BC051676
 VERSION BC051676.1 GI:30354527
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2431)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullany, S.J., Bosak, S.A., McGowan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Pahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalske, U., Smailus, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 JOURNAL 12477932
 MEDLINE 22388257
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 2431)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (30-APR-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
 USA
 REMARK
 COMMENT NTH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Minoru Ko
 CDNA Library Preparation: Yulan Piao and Minoru Ko
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisse, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 111 Row: i Column: 20
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 23943897.
 Location/Qualifiers
 1. 2431
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="MGC:60994 IMAGE:30017484"
 /issue_type="Embryo, whole, mouse, 7.5-dpc"


```

variation      3115. .3129
                /gene="IRAK4"
                /frequency="0.72"
                /replace="n"
variation      3189
                /gene="IRAK4"
                /frequency="0.01"
                /replace="g"
variation      3536
                /gene="IRAK4"
                /frequency="0.14"
                /replace="c"
variation      3573
                /gene="IRAK4"
                /frequency="0.01"
                /replace="t"
variation      3819
                /gene="IRAK4"
                /frequency="0.17"
                /replace="t"
variation      3877
                /gene="IRAK4"
                /frequency="0.01"
                /replace="g"
variation      4144
                /gene="IRAK4"
                /frequency="0.04"
                /replace="a"
variation      4438. .4610
                /rpt_family="MIR"
                /rpt_type=dispersed
variation      4602
                /gene="IRAK4"
                /frequency="0.01"
                /replace="t"
variation      4797
                /gene="IRAK4"
                /frequency="0.02"
                /replace="a"
variation      4972
                /gene="IRAK4"
                /frequency="0.31"
                /replace="n"
variation      5759. .5797
                /rpt_family="I2"
                /rpt_type=dispersed
variation      5821
                /gene="IRAK4"
                /frequency="0.02"
                /replace="g"
variation      5880
                /gene="IRAK4"
                /frequency="0.01"
                /replace="c"
variation      5954
                /gene="IRAK4"
                /frequency="0.01"
                /replace="t"
variation      5964. .6036
                /rpt_family="MER1_type"
                /rpt_type=dispersed
variation      6023. .6590
                /rpt_family="MER1_type"
                /rpt_type=dispersed
variation      6204
                /gene="IRAK4"
                /frequency="0.04"
                /replace="t"
variation      6661. .6662
                /gene="IRAK4"
                /frequency="0.04"
                /replace="n"
variation      6998

```

```

                /gene="IRAK4"
                /frequency="0.04"
                /replace="g"
repeat_region 7015. .7073
                /rpt_family="Mariner"
                /rpt_type=dispersed
repeat_region 7074. .7366
                /rpt_family="Alu"
                /rpt_type=dispersed
variation      7122
                /gene="IRAK4"
                /frequency="0.01"
                /replace="a"
variation      7147
                /gene="IRAK4"
                /frequency="0.01"
                /replace="a"
variation      7234
                /gene="IRAK4"
                /frequency="0.04"
                /replace="c"
variation      7367. .8636
                /rpt_family="Mariner"
                /rpt_type=dispersed
variation      7392
                /gene="IRAK4"
                /frequency="0.01"
                /replace="c"
variation      7597
                /gene="IRAK4"
                /frequency="0.01"
                /replace="c"
variation      7794
                /gene="IRAK4"
                /frequency="0.09"
                /replace="g"

```

Query Match 51.6%; Score 151.8; DB 9; Length 33033;

Best Local Similarity 95.7%; Pred. No. 1.6e-28; Matches 156; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY      123 TCAGTTTCACATAGAGAGATTGAGCATTACTCAAACTGAAAAAGTCCCACTTCTGA 182
         |||||
Db      14142 TTACTTACTTTNAGAGAGATTGAGCATTACTCAAACTGAAAAAGTCCCACTTCTGA 14201
         |||||
QY      183 ATTACTGTTTGACTGGGGGACCAACAATTGCACAGTTGATCTTGTGATCTTTTGAT 242
         |||||
Db      14202 ATTACTGTTTGACTGGGGGACCAACAATTGCACAGTTGATCTTGTGATCTTTTGAT 14261
         |||||
QY      243 CCAAAATGAATTTTGTCTCCGCGAGTCTTTTGTCCCAAGT 285
         |||||
Db      14262 CCAAAATGAATTTTGTCTCCGCGAGTCTTTTGTCCCAAGT 14304
         |||||

```

```

RESULT 15
AC093012      118572 bp      DNA      linear      PRI 24-SEP-2002
LOCUS
DEFINITION
Human BAC Library) complete sequence.
AC093012
VERSION
AC093012.9 GI:21629105
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
AUTHORS
Muzny,D.M., Adams,C., Adio-Oduola,B., All-ouman,F.R., Allen,C.,
Alstbrooks,S.L., Amaratunga,H.C., Are,J.R., Ayala,M., Banks,T.,
Barbarta,J., Benton,J., Bimarge,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,

```

	Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dedertich, D.A., Delaney, K.L., Delgado, O., Benn, A.U., Ding, Y., Dimin, H.J., Duttaiwale, K.R., Draper, H., Dugan-Rocha, S., Durbin, K.H., Barnhart, K.C., Edgar, D., Edwards, C.C., Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Franz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, U.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, U., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Hollway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Huliy, S., Home, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvan, J., Kovac, C., Kratochvic, U., Kuresh, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, U., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W., Louisgeed, H., Lozano, R.U., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Mondrel, I., Martin, R., Martindale, A., Martinez, E., Massey, E., McWhiney, E., Meleod, M.P., Meador, M., Mel, G., Mescher, S., Metzger, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabhat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokhekwo, S., Oguh, M., Okunodu, G., Orangunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pul, L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., RojudoKan, I., Rolfe, M., Ruiz, S., Savary, G., Schetter, S., Scott, G., Shen, H., Shim, C., Shooshtarian, N., Slisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabors, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telirod, B., Thomas, N., Thomas, S., Usmani, K., Vaquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wleczyk, R., Woodem, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Kucheralapati, R., Weinstein, G., and Gibbs, R.
TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 118572)
AUTHORS	Morley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (09-AUG-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 118572)
AUTHORS	Morley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (04-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	4 (bases 1 to 118572)
AUTHORS	Morley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (21-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	5 (bases 1 to 118572)
AUTHORS	Morley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (29-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	6 (bases 1 to 118572)
AUTHORS	Morley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (03-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	7 (bases 1 to 118572)
AUTHORS	Morley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (24-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One

COMMENT

Baylor Plaza, Houston, TX 77030, USA
On Jun 29, 2002 this sequence version replaced gi:21535906
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions for the remainder of sequenced and submitted once, so the sequence for the adjacent clones. The insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smrt and P. Green, unpublished) for Human and Mouse sequences.

Genes and Region of consensus similarity are identified by BLAST (Altschul et al., 1990) using the NCBI database (NCBI, 2002). Genes demonstrating a similarity (expect < 1e-34) to EST and cDNA sequences. Genes demonstrate that maintained sequence flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.hgsc.bcm.tmc.edu/8088/quality.info/genbank.annotation.html>.

FEATURES

Source

```
misc_feature
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="12"
      /clone="RP11-210N13"
      complement(1..2004)
      /note="overlaps bases 1..2004 of clone AC016143"
      /function="Clone overlap"
repeat_region
      869..903
      /rpt_family="HSMAR1"
repeat_region
      1047..1467
      /rpt_family="MLTK"
repeat_region
      1644..1760
      /rpf_family="MIR"
repeat_region
      1985..2295
      /rpf_family="AluY"
repeat_region
      2296..2325
      /rpf_family="AT_rich"
repeat_region
      complement(2730..3024)
      /rpf_family="AluSg"
repeat_region
      4092..4211
      /rpf_family="HERA"
repeat_region
      4389..4415
      /rpf_family="AT_rich"
repeat_region
      4502..4561
      /rpf_family="AT_rich"
repeat_region
      complement(4612..4778)
      /rpf_family="MIR"
repeat_region
      5859..5926
      /rpf_family="AT_rich"
repeat_region
      complement(6385..6636)
      /rpf_family="AluSg"
```

```
repeat_region complement(7252, .7562)
/rpc_family="L2"
repeat_region 7574, .8066
/rpc_family="MLTIF"
repeat_region complement(8266, .8438)
/rpc_family="MIR"
repeat_region 8889, .8927
/rpc_family="AT_rich"
repeat_region 10113, .10252
/rpc_family="L1M4"
repeat_region 10433, .10726
/rpc_family="AluY"
repeat_region 11181, .11552
/rpc_family="THE1C"
repeat_region complement(11739, .11889)
/rpc_family="MERSB"
repeat_region 11910, .12124
/rpc_family="MERS8A"
repeat_region 12249, .12617
/rpc_family="THE1C"
repeat_region 14290, .14310
/rpc_family="AT_rich"
repeat_region 14363, .14670
/rpc_family="AluSq"
repeat_region complement(14742, .15039)
/rpc_family="AluSx"
repeat_region 15669, .15691
/rpc_family="(TTG)n"
repeat_region complement(15692, .15972)
/rpc_family="AluSc"
repeat_region 16265, .16352
/rpc_family="L1PA16"
repeat_region 16353, .16631
/rpc_family="AluO"
repeat_region 16632, .17007
/rpc_family="L1PA16"
repeat_region 17329, .17356
```

Query Match 51.6%; Score 151.8; DB 9; Length 118572;

Best Local Similarity 95.7%; Pred. No. 1.4e-28;

Matches 156; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY 123 TCAGTTTCACATAGAGATTGAGATTACTTCAAACTGAAAAAGTCCACTCTGA 182
DB 38156 TTACTTACTTTTAGAGATTGAGATTCTTCAAACTGAAAAAGTCCACTCTGA 38215
QY 183 ATTACTGTTTACTGGGACACCAAAATGACAGTTGGTGTATCTTGATCTTTTGA 242
DB 38216 ATTACTGTTTACTGGGACACCAAAATGACAGTTGGTGTATCTTGATCTTTTGA 38275
QY 243 CCAAAATGAATTTTGTCTGCGAGTCTTTGTCTCCAGAT 285
DB 38276 CCAAAATGAATTTTGTCTGCGAGTCTTTGTCTCCAGAT 38318
```

Search completed: January 18, 2004, 00:16:05
job time : 1518 secs

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 23:50:52 ; Search time 51 Seconds
(without alignment)
2544.444 Million cell updates/sec

Title: US-10-001-254-5
Perfect score: 294
Sequence: 1 acatattgctgctgctcctcaaa.....tgcctccagatgcttccc 294

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	294	100.0	833	4 US-09-166-350-10	Sequence 10, Appl
2	40.4	13.7	1879	4 US-09-601-091-1	Sequence 1, Appl
3	40.4	13.7	2024	4 US-09-601-091-3	Sequence 3, Appl
4	40.4	13.7	2024	4 US-09-398-395A-51	Sequence 51, Appl
5	40.4	13.7	2024	4 US-09-887-586A-51	Sequence 51, Appl
6	40.4	13.7	2024	4 US-09-885-752-51	Sequence 51, Appl
7	40.4	13.7	2024	4 US-09-903-012B-51	Sequence 51, Appl
8	33.8	11.5	580073	4 US-08-545-528D-1	Sequence 1, Appl
9	33.6	11.4	2393	4 US-08-961-527-90	Sequence 90, Appl
10	32.8	11.4	202001	4 US-09-734-674-3	Sequence 3, Appl
11	32.6	11.1	1728	4 US-09-036-731A-1	Sequence 1, Appl
12	32.4	11.0	98844	4 US-09-791-211-10	Sequence 10, Appl
13	32.2	11.0	4402	4 US-09-484-970B-135	Sequence 135, Appl
14	31.8	10.9	2274	4 US-09-107-532A-3535	Sequence 3535, Ap
15	31.8	10.8	1557	4 US-09-134-001C-1614	Sequence 1614, Ap
16	31.8	10.8	2507	4 US-09-620-312D-120	Sequence 120, App
17	31.8	10.8	1664976	4 US-08-916-421B-1	Sequence 1, Appl
18	31.4	10.7	12720	1 US-08-403-866-11	Sequence 11, Appl
19	31.1	10.5	2126	2 US-08-545-745-1	Sequence 1, Appl
20	30.8	10.5	1550	4 US-09-427-501-1	Sequence 1, Appl
21	30.8	10.5	15016	4 US-09-601-198-60	Sequence 60, Appl
22	30.8	10.5	319608	4 US-09-539-333D-1	Sequence 1, Appl
23	30.8	10.5	319608	4 US-09-679-409-1	Sequence 1, Appl
24	30.8	10.5	1230025	4 US-09-198-452A-1	Sequence 1, Appl
25	30.6	10.4	1479	4 US-09-328-352-897	Sequence 897, Appl
26	30.4	10.3	1362	3 US-08-929-329-4	Sequence 4, Appl
27	30.2	10.3	2254	2 US-08-635-066-1	Sequence 1, Appl

C 28	30	10.2	195	1	US-08-324-243-20	Sequence 20, Appl
C 29	30	10.2	195	1	US-08-532-390-20	Sequence 20, Appl
C 30	30	10.2	195	3	US-08-717-294-20	Sequence 20, Appl
C 31	30	10.2	195	5	PCT-US95-11511-20	Sequence 20, Appl
C 32	30	10.2	1632	1	US-08-324-243-34	Sequence 34, Appl
C 33	30	10.2	1632	1	US-08-532-390-34	Sequence 34, Appl
C 34	30	10.2	1632	3	US-08-717-294-34	Sequence 34, Appl
C 35	30	10.2	1632	5	PCT-US95-11511-34	Sequence 34, Appl
C 36	30	10.2	2481	1	US-08-324-243-35	Sequence 35, Appl
C 37	30	10.2	2481	1	US-08-532-390-35	Sequence 35, Appl
C 38	30	10.2	2481	3	US-08-717-294-35	Sequence 35, Appl
C 39	30	10.2	2481	5	PCT-US95-11511-35	Sequence 35, Appl
C 40	30	10.2	2571	4	US-09-552-950-4	Sequence 4, Appl
C 41	30	10.2	11770	4	US-08-961-527-172	Sequence 172, App
C 42	29.8	10.1	4648	4	US-09-207-914-24	Sequence 24, Appl
C 43	29.8	10.1	4655	4	US-09-207-914-20	Sequence 20, Appl
C 44	29.6	10.1	1356	1	US-08-592-936B-22	Sequence 22, Appl
C 45	29.6	10.1	1356	1	US-08-788-928A-2	Sequence 2, Appl

ALIGNMENTS

DNA of Seq ID#5

RESULT 1
US-09-166-350-10
Application US/09166350A
Patent No. 6440663
GENERAL INFORMATION:
APPLICANT: Scanlan, Matthew
APPLICANT: Chen, Yao
APPLICANT: Stockert, Elisabeth
APPLICANT: Old, Lloyd
APPLICANT: Jager, Elke
APPLICANT: Knuth, Alex
TITLE OF INVENTION: Renal Cancer Associated Antigens and
FILE REFERENCE: 10461/7051
CURRENT APPLICATION NUMBER: US/09/166,350A
EARLIER FILING DATE: 1998-10-05
EARLIER APPLICATION NUMBER: US 09/166,350
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 833
TYPE: DNA
ORGANISM: Homo sapiens
US-09-166-350-10

Query Match 100.0%; Score 294; DB 4; Length 833;
Best Local Similarity 100.0%; Pred. No. 3.4e-82;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ACATATGTCGCTGCTCCTCATATGTTGACTAATATAGGAACTGTCAGATTATATGATCCT	60
DB	74	ACATATGTCGCTGCTCCTCATATGTTGACTAATATAGGAACTGTCAGATTATATGATCCT	133
QY	61	CAAGAGATGGAAGAGTTAGCTGATTAATTAACATCTGATGATAGATAC	120
DB	134	CAAGAGATGGAAGAGTTAGCTGATTAATTAACATCTGATGATAGATAC	193
QY	121	AATCATGTTCACTAATGAGATTTGAGCACTTCAATCGAAGAAAGTCCACTTCT	180
DB	194	AATCATGTTCACTAATGAGATTTGAGCACTTCAATCGAAGAAAGTCCACTTCT	253
QY	181	GAATTACTGTTGACTGCGGACCAAAATTCACAGTTGATCTTGTGATCTTTTG	240
DB	254	GAATTACTGTTGACTGCGGACCAAAATTCACAGTTGATCTTGTGATCTTTTG	313
QY	241	ATTCAAATGAATTTTCTCTGAGATCTTTTGTCCAGATGCTGTTCC	294
DB	314	ATTCAAATGAATTTTCTCTGAGATCTTTTGTCCAGATGCTGTTCC	367

RESULT 2

US-09-601-091-1
; Sequence 1, Application US/09601091
; Patent No. 6342380
; GENERAL INFORMATION:
; APPLICANT: Colby, S. M. et al.
; TITLE OF INVENTION: Germacrene C Synthase Gene of Lycopersicon Esculentum
; FILE REFERENCE: 4630-55758
; CURRENT APPLICATION NUMBER: US/09/601,091
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: PCT/US99/02133
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073,579
; PRIOR FILING DATE: 1998-02-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1879
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1685)
US-09-601-091-1

Query Match

Best Local Similarity 13.7%; Score 40.4; DB 4; Length 1879;
Matches 98; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 45 AGATTTTATGATCTCTCAAGAGATGAGAGAGATTAGCTGTATTAATAAACCATC 104
DB 761 AGAGCTTAGATGATCTTACAAAGGTGGTGAATAATTTGCAATTAATATCCATA 820
QY 105 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 164
DB 821 TGCAGAGACAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 880
QY 165 AAAAGTCCCATCTTGAATTAATGATGATGATGATGATGATGATGATGATGATG 224
DB 881 ATATAGTCGTGAG 940
QY 225 TCTTGTGATCTTT 238
DB 941 CACTTTGATGCTT 954

RESULT 3

US-09-601-091-3
; Sequence 3, Application US/09601091
; Patent No. 6342380
; GENERAL INFORMATION:
; APPLICANT: Colby, S. M. et al.
; TITLE OF INVENTION: Germacrene C Synthase Gene of Lycopersicon Esculentum
; FILE REFERENCE: 4630-55758
; CURRENT APPLICATION NUMBER: US/09/601,091
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: PCT/US99/02133
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073,579
; PRIOR FILING DATE: 1998-02-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)..(1678)
US-09-601-091-3

Query Match

13.7%; Score 40.4; DB 4; Length 2024;

Best Local Similarity 50.5%; Pred. No. 0.0043;
Matches 98; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 45 AGATTTTATGATCTCTCAAGAGATGAGAGAGATTAGCTGTATTAATAAACCATC 104
DB 754 AGAGCTTAGATGATCTTACAAAGGTGGTGAATAATTTGCAATTAATATCCATA 813
QY 105 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 164
DB 814 TGCAGAGACAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 873
QY 165 AAAAGTCCCATCTTGAATTAATGATGATGATGATGATGATGATGATGATGATG 224
DB 874 ATATAGTCGTGAG 933
QY 225 TCTTGTGATCTTT 238
DB 934 CACTTTGATGCTT 947

RESULT 4

US-09-398-395A-51
; Sequence 51, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64687721, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)..(1675)
; OTHER INFORMATION: VFMT germacrene C synthase
US-09-398-395A-51

Query Match

Best Local Similarity 13.7%; Score 40.4; DB 4; Length 2024;
Matches 98; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 45 AGATTTTATGATCTCTCAAGAGATGAGAGAGATTAGCTGTATTAATAAACCATC 104
DB 754 AGAGCTTAGATGATCTTACAAAGGTGGTGAATAATTTGCAATTAATATCCATA 813
QY 105 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 164
DB 814 TGCAGAGACAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 873
QY 165 AAAAGTCCCATCTTGAATTAATGATGATGATGATGATGATGATGATGATGATG 224
DB 874 ATATAGTCGTGAG 933
QY 225 TCTTGTGATCTTT 238
DB 934 CACTTTGATGCTT 947

RESULT 5

```
US-09-887-586A-51
; Sequence 51, Application US/09887586A
; Patent No. 6495354
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64953541, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VFNT germacrene C synthase
US-09-887-586A-51
```

```
Query Match      13.7%; Score 40.4; DB 4; Length 2024;
Best Local Similarity 50.5%; Pred. No. 0.0043;
Matches 98; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
```

```
QY 45 AGATTATTGATCTCTCAAGAGATGAGAAAGTTAGCTGTACTTATTAATAAACATC 104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 754 AGAGCTTAGTGATCTTCAAGGTGTGAAAGATTGGATTTCGAATAATATATCCATA 813
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 105 TGGTGATGATGATACATCATGTTTCATAGAGATTGAGACATTAAGCATTAAGTCAACTGG 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 814 TGCAGAGACAGGTTGGTTGATGATCTTCTGTGATATAGAGTGTATTTTGACCAAA 873
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 165 AAAAAGTCCCACTTCTGAATTACTGTTTGAAGTGGGACCAAAATTCACAGTTGGTGA 224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 874 ATATAGTCGTGCGGAGAAAATGATGACAAAAGTACTCACTGACCTCCATTATGACGA 933
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 225 TCTTGGATCTTT 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 934 CACTTTTGATGCTT 947
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 6
US-09-895-752-51
; Sequence 51, Application US/09895752
; Patent No. 6559297
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65592971, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/895,752
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VFNT germacrene C synthase
US-09-895-752-51
```

```
Query Match      13.7%; Score 40.4; DB 4; Length 2024;
Best Local Similarity 50.5%; Pred. No. 0.0043;
Matches 98; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
```

```
QY 45 AGATTATTGATCTCTCAAGAGATGAGAAAGTTAGCTGTACTTATTAATAAACATC 104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 754 AGAGCTTAGTGATCTTCAAGGTGTGAAAGATTGGATTTCGAATAATATATCCATA 813
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 105 TGGTGATGATGATACATCATGTTTCATAGAGATTGAGACATTAAGCATTAAGTCAACTGG 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 814 TGCAGAGACAGGTTGGTTGATGATCTTCTGTGATATAGAGTGTATTTTGACCAAA 873
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 165 AAAAAGTCCCACTTCTGAATTACTGTTTGAAGTGGGACCAAAATTCACAGTTGGTGA 224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 874 ATATAGTCGTGCGGAGAAAATGATGACAAAAGTACTCACTGACCTCCATTATGACGA 933
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 225 TCTTGGATCTTT 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 934 CACTTTTGATGCTT 947
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 7
US-09-903-012B-51
; Sequence 51, Application US/09903012B
; Patent No. 6569656
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65696561, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012B
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VFNT germacrene C synthase
US-09-903-012B-51
```

```
Query Match      13.7%; Score 40.4; DB 4; Length 2024;
Best Local Similarity 50.5%; Pred. No. 0.0043;
Matches 98; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
```

```
QY 45 AGATTATTGATCTCTCAAGAGATGAGAAAGTTAGCTGTACTTATTAATAAACATC 104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 754 AGAGCTTAGTGATCTTCAAGGTGTGAAAGATTGGATTTCGAATAATATATCCATA 813
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 105 TGGTGATGATGATACATCATGTTTCATAGAGATTGAGACATTAAGCATTAAGTCAACTGG 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

1 APPLICATION NUMBER: US/08/961,527
2 FILING DATE:
3 CLASSIFICATION: 424
4 PRIOR APPLICATION DATA:
5 APPLICATION NUMBER:
6 FILING DATE:
7 ATTORNEY/AGENT INFORMATION:
8 NAME: Brookes, A. Anders
9 REGISTRATION NUMBER: 36,373
10 REFERENCE/DOCKET NUMBER: PB340P1
11 TELECOMMUNICATION INFORMATION:
12 TELEPHONE: (301) 309-8504
13 TELEFAX: (301) 309-8512
14 INFORMATION FOR SEQ ID NO: 90:
15 SEQUENCE CHARACTERISTICS:
16 LENGTH: 2393 base pairs
17 TYPE: nucleic acid
18 STRANDEDNESS: double
19 TOPOLOGY: linear
20
21 US-08-961-527-90
22
23 Query Match 11.4%; Score 33.6; DB 4; Length 2393;
24 Best Local Similarity 48.9%; Pred. No. 0.6;
25 Matches 99; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
26
27 QY 37 AACGTGTCAGATTATTGATCTCTCAAGAGATGGAAGAGTGTAGCTGATTTAA 96
28 Db 1113 AACCTTTTGTAGTACCAACGAACTAAAGAACTCGCAACAAAGAGATATTCCTATTAA 1054
29
30 QY 97 AAACCATCGGTGATATATGATATCATCATGATTTCACATAGAGAGATTGAGATTACTT 156
31 Db 1053 ATAGGAAGTATGATGATGTGGCAATTACTGATTTTGAATTAATGATATGTTAGCAATGATT 994
32
33 QY 157 CAACCTGAAAAAGTCCCACTTCTGATTAATGATGTTGACTGGGGCCACCAAAATTGCACA 216
34 Db 993 GTTACTCTAAAGAAATATCTGCTTAATTTGGCAAAATTAATTGAGAGCTAACAACCTTGACA 934
35
36 QY 217 GTTG 220
37 Db 933 AATG 930
38
39 RESULT 10
40 US-09-734-674-3/c
41 Sequence 3, Application US/09734674
42 Patent No. 6498022
43 GENERAL INFORMATION:
44 APPLICANT: WEI, Ming-Hui et al
45 TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
46 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS
47 FILE REFERENCE: C1001018
48 CURRENT APPLICATION NUMBER: US/09/734,674
49 CURRENT FILING DATE: 2000-12-13
50 NUMBER OF SEQ ID NOS: 4
51 SOFTWARE: FastSeq for Windows Version 4.0
52 SEQ ID NO 3
53 LENGTH: 202001
54 TYPE: DNA
55 ORGANISM: Human
56 FEATURE:
57 NAME/KEY: misc_feature
58 LOCATION: (1)..(202001)
59 OTHER INFORMATION: n = A,T,C or G
60
61 US-09-734-674-3
62
63 Query Match 11.2%; Score 32.8; DB 4; Length 202001;
64 Best Local Similarity 58.0%; Pred. No. 6.2;
65 Matches 58; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
66
67 QY 62 AAGAAGATGGAAGAGTTAGCTGTGCTATTTAAAAAACCATCTGGTGTAGTAGTTACA 121
68 Db 91622 AAGAAGGTATCAAAAAATTAGCTGTATCTATTAATAATATGCTAAGGGGGAAGATATACC 91563

```

QY	122	ATCAGTTTCACATAAGGAGATTGAAGCACTACTCAAC	161
Db	91562	TGCAATTTCTTCTAAAGTGTGGCATGCATGTCACGCAAC	91523

RESULT 11

```

US-09-096-731A-1
Sequence 1, Application US/09096731A
Patent No. 6489454
GENERAL INFORMATION:
APPLICANT: Liu, Olong
APPLICANT: Hengartner, Michael O.
APPLICANT: Bogaert, Thierry Andre Oliver Eddy
APPLICANT: Van Criekeinge, Wim Maria Rene
TITLE OF INVENTION: ENGLIMPLEM GENE AND USES THEREOF
FILE REFERENCE: CSH:97-075A
CURRENT APPLICATION NUMBER: US/09/096,731A
CURRENT FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: US 60/072,324
PRIOR FILING DATE: 1998-01-23
NUMBER OF SEQ. ID NOS: 44
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO. 1
LENGTH: 1728
TYPE: DNA
ORGANISM: Caenorhabditis elegans
FEATURE:
NAME/KEY: CDS
LOCATION: (22) ... (1500)
US-09-096-731A-1

```

Query Match	Score	DB	Length
11.1%	32.6	4	1728

OY	46	GATTTTATGATCCCTCAGAGAAGATTGAAGAAGTAGCTGTAGCATTAATAAAAAACATC	105
Db	1132	GATTCGTTGATCCCAGACTGGAGAAAANAAGCACATCGCAAGCTGAGTAAATTCATTCC	1190

Oy	106	GGGATGATGATAGATCAATCAGTTTCACATPAGGAGATTGAGCACTTACTTCAACTGGA	165
Db	1192	GGTGGCGGACTTCCTCAGTGGCATTCCMAATGGTAAAGAACCAACCACATAGCCTCGCT	1251

QY	166	AAAGTCCCACTTCTGAATTACT	188
Db	1252	GAAGTCTCGCTTCTGAAGCAAT	1274

RESULT 12

```

US-09-791-211-10/c
Sequence 10, Application US/09791211
Patent No. 6448080
GENERAL INFORMATION:
APPLICANT: Donna T. Ward
TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
FILE REFERENCE: RTS-0205
CURRENT APPLICATION NUMBER: US/09/791,211
CURRENT FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 90
SEQ ID NO 10
LENGTH: 98844
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: 24362
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 64383
OTHER INFORMATION: unknown
NAME/KEY: unsure

```

```

? LOCATION: 65468
? OTHER INFORMATION: unknown
? NAME/KEY: unsure
? LOCATION: 65469
? OTHER INFORMATION: unknown
? NAME/KEY: unsure
? LOCATION: 65470
? OTHER INFORMATION: unknown
? NAME/KEY: unsure
? LOCATION: 65471
? OTHER INFORMATION: unknown
? NAME/KEY: unsure
? LOCATION: 87110
? OTHER INFORMATION: unknown
? NAME/KEY: unsure
? LOCATION: 89049
? OTHER INFORMATION: unknown
? OTHER INFORMATION: unknown
? OTHER INFORMATION: unknown
IS-09-791-211-10
```

Query March	11.0%	Score 32.4;	DB 4;	Length 98844;
Best Local Similarity	53.1%;	Pred. No.6.2;		
Matches 69;	Conservative 0;	Mismatches 61;	Indels 0;	Gaps 0;

QY	Db	QY	Db
51	7415	111	7355
TATTGATCCCTCAGAAAGATGGAAAGATTAGCTGAGCTTTTAAAAAACATCTGGGGA	TATATATTTTTCAGTATTAATGGAAAAAATCCGTACTGTTATTTATTTGACATGCTATTT	TGATAGTACATACAGTTTCACATTAAGAGATTTGAAGCTTACTTCAAACTGGAAAAAG	TAAATCTTTTAAATTTATTTTGGTAATGAGATTTTAAATCATATTTTGGAAATGGGAAA
110	7355	170	7296

QY	171	TCCCACTTCT	180
Db	7295	ACATACATAT	7286

RESULT 13

```

US-09-484-970B-135
Sequence 135, Application US/09484970B
Patent No. 6426186
GENERAL INFORMATION:
APPLICANT: Jones, Karen A.
APPLICANT: Volkmuth, Wayne
APPLICANT: Walker, Michael G.
TITLE OF INVENTION: BONE REMODELING GENES
FILE REFERENCE: PB-0014 US
CURRENT APPLICATION NUMBER: US/09/484,970B
CURRENT FILING DATE: 2000-01-18
NUMBER OF SEQ ID NOS: 172
SOFTWARE: PERL Program
SEQ ID NO 135
LENGTH: 4402
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6426186 252875 (2874502CB1)
US-09-484-970B-135

```

Query Match	11.0%	Score 32.2;	DB 4;	Length 4402;
Similarity	51.9%	Pred. No.2.1;		
Best Local	0;	Mismatches	88;	Indels 2;
Matches 97;	Conservative			Gaps 1;

QY 110 TGTATGATCAATCACTGTTTCACATAGAGATTGTAAGCATTACTTCAACTGGAATAA 165
Db 2771 ATTAACATATGAATTCAGTTAAAGAATGTGTTCATTAATGACTTTTAACTGCTAATAA 2830

Qy 170 GTCCCACTT--CTGAATTACTCTTTGACTGGGGACCAACAATTGCAAGTGGTGACT 227

Db 2831 ATATTACTTGACGAAGTACTTGATGTATGTTATCTGAAATTCGGAGTATTTGGTCT 2890

228 TGTGATCTTTGATCCAAATGAATTTTGTCTCCTGCGAGTCTTTGTCTCCAGATGC 287

Db 2891 GTTCTTGTCTAAATAAGTCTGTTTGTCTAGTCTTCAGAAATATTTATTTCTGAGA 2950
QY 288 TGTTCCTC 294
Db 2951 TGTTCCTC 2957

RESULT 14

US-09-107-532A-3535
; Sequence 3535, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-8277
; TELEFAX: (781)893-8207
; INFORMATION FOR SEQ ID NO: 3535:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2274 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...2274
; SEQUENCE DESCRIPTION: SEQ ID NO: 3535:
US-09-107-532A-3535

Query March 10.9%; Score 32; DB 4; Length 2274;
Best Local Similarity 52.2%; Pred. No. 1.9;
Matches 71; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 30 AATTGAGAGCTGTCAAGTTTATTTGATCTCAAGAGATGAGAGAGTTAGCTGTAGC 89
Db 1559 AATCACTTAGCGCGCAATTTAGATTAGTGAATGAACAAACAGACGTACCAA 1618
QY 90 TATTAAAAACCATCTGTGTATGATAGATCAATCAGTTTCACCTAAGAGATTGAGC 149
Db 1619 AAGGAAATACGAAAGGTTCAGCGGTTCAATGCAAGCAAGTAGCATATGGATTGAGAA 1678

QY 150 ATTACTCAAACTGGA 165
Db 1679 ATTTCACAAAAAGAA 1694

RESULT 15

US-09-134-001C-1614
; Sequence 1614, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1614
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1614

Query Match 10.8%; Score 31.8; DB 4; Length 1557;
Best Local Similarity 48.6%; Pred. No. 1.9;
Matches 87; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 87 AGCTATTAAAAAACCATCTGTGTGATGATGATACATCAGTTTCACATAGAGATTGGA 146
Db 1092 AGTAATTAATCAAAAACAATTGATTAATCAAAAATTTATTTGAATGTGAAAAAGA 1151
QY 147 AGCATTAATCAAACTGGAAGAAAGTCCACTTCTGAATTACTGTTGACTGGGACCAAC 206
Db 1152 AGGCAAACTAGAGACTGTGTGGAACAGATGATTTCTACCGGTATTCTTGTGAACCAAC 1211
QY 207 AATTGCACAGTTGTGTGATCTTGTGATCTTTGATCCAAATGAAATTTTGTCTCTG 265
Db 1212 GATTTTCTCGGACTACATCTGCGGATGCTATCATGCAAGAAATTTTGGACCAAG 1270

Search completed: January 18, 2004, 00:40:18
Job time : 55 secs